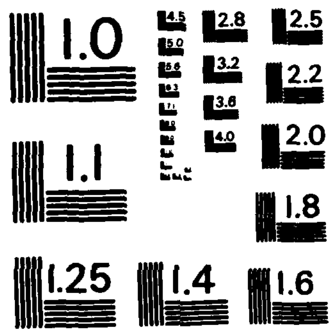
AD-A152 083 PROCESSES INVOLVED IN WRITING EFFECTIVE PROCEDURAL 1/2
INSTRUCTIONS(U) AMERICAN INSTITUTES FOR RESEARCH
WASHINGTON DC V M HOLLAND ET AL. 28 FEB 85
UNCLASSIFIED N00014-83-C-0590 F/G 5/7 NL


1.0
4.5
5.0
5.6
6.3
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.1
1.8
1.25
1.4
1.6

2

Processes Involved in Writing Effective Procedural Instructions

Final Report

Melissa Holland
Andrew Rose
Robin Dean
Stephen Dory

February 28, 1985

The work reported here was sponsored by the Personnel and Training Research Programs. Psychological Sciences Division, Office of Naval Research, under Contract No. N00014-83-C-0590, Contract Authority Identification Number, NR154-525. Approved for public release; distribution unlimited. Reproduction in whole or part is permitted for any purpose of the United States Government.

AMERICAN INSTITUTES FOR RESEARCH/1055 Thomas Jefferson Street, NW, Washington, DC 20007

REPORT DOCUMENTATION PAGE		*READ INSTRUCTIONS BEFORE COMPLETING FORM*
1. REPORT NUMBER Technical Report No. 1	2. GOVT ACCESSION NO. AD-A152-083	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Processes Involved in Writing Effective Procedural Instructions		5. TYPE OF REPORT & PERIOD COVERED Final Report (3 Aug 83 - 31 Dec 84)
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) V. Melissa Holland Andrew M. Rose Robin A. Dean Steven L. Dory		8. CONTRACT OR GRANT NUMBER(s) N00014-83-C-0590
9. PERFORMING ORGANIZATION NAME AND ADDRESS American Institutes for Research 1055 Thomas Jefferson St., N.W. Washington, D.C. 20007		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61153N 42 RR04206;RR0420601 NR154-525
11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Program Office of Naval Research (Code 442PT) Arlington, VA 22217		12. REPORT DATE 28 Feb 85
		13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

(U) Authoring; (U) Discourse; (U) Technical Documentation; (U) Linguistics (U) Psycholinguistics; (U) Technical Writing; (U) Text

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This study examined instruction writing, seeking to distinguish effective from ineffective writers in terms of the processes they use. The study involved four phases: (1) observing 16 writers composing a set of instructions for one of two procedural tasks--item assembly and knot-tying; (2) identifying effective instructions through user testing; (3) observing writers revising their own instructions, first before, then after viewing a videotaped user; (4) identifying effective revisions through user testing.

Approved for public release; distribution unlimited

Analyses were conducted (1) of selected text features of the instructions and (2) of various dimensions of the writing process, including: surface measures of writers' behavior, such as time spent reviewing; records of the number and kinds of textual changes writers made; protocols of the writers verbalizing their thoughts.

Analyses of text features showed that the primary distinction between effective and ineffective instructions lay in the sufficiency of the information they contained to convey each task step. Sentence and word variables, formatting, and the degree of text hierarchy were not consistently distinguishing.

Analyses of writers' surface behaviors and change patterns, although they revealed striking variation, showed little that consistently distinguished effective from ineffective writers. Information planning in assembly and whole-text review in knot-tying characterized some effective writers.

Analyses of protocols in terms of the types of problems writers defined showed that the primary distinction between effective and ineffective writers lay in how they dealt with selecting information from their base of procedural task knowledge. Effective writers tended to define more problems related to selecting information (e.g., assessing informational sufficiency, level of detail, need for feedback). They also tended to test the adequacy of information by anticipating users' responses at choice points in their emerging texts. Ineffective writers either failed to consider problems of selecting information or considered them but applied predetermined evaluation criteria. Although these criteria were often user-based, they lacked the flexibility and specificity of constructing the response of imagined users to particular alternatives. Results are discussed in terms of instruction writers' ability to represent and balance conflicts between competing sets of criteria, such as being informative and being succinct.

Results of the revising phase showed that watching a user helped all writers detect informational ambiguities in their instructions.

Processes Involved in Writing Effective Procedural Instructions

Final Report

Melissa Holland
Andrew Rose
Robin Dean
Stephen Dory

February 28, 1985

The work reported here was sponsored by the Personnel and Training Research Programs. Psychological Sciences Division, Office of Naval Research, under Contract No. N00014-83-C-0590, Contract Authority Identification Number, NR154-525. Approved for public release; distribution unlimited. Reproduction in whole or part is permitted for any purpose of the United States Government.

AMERICAN INSTITUTES FOR RESEARCH/1055 Thomas Jefferson Street, NW, Washington, DC 20007

Abstract

This study examined instruction writing, seeking to distinguish effective from ineffective writers in terms of the processes they use. The study involved four phases: (1) observing 16 writers composing a set of instructions for one of two procedural tasks--item assembly and knot-tying; (2) identifying effective instructions through user-testing; (3) observing writers revising their own instructions, first before, then after viewing a videotaped user; (4) identifying effective revisions through user testing.

Analyses were conducted (1) of selected text features of the instructions and (2) of various dimensions of the writing process, including: surface measures of writers' behavior, such as time spent reviewing; records of the number and kinds of textual changes writers made; protocols of the writers verbalizing their thoughts.

Analyses of text features showed that the primary distinction between effective and ineffective instructions lay in the sufficiency of the information they contained to convey each task step. Sentence and word variables, formatting, and the degree of text hierarchy were not consistently distinguishing.

Analyses of writers' surface behaviors and change patterns, although they revealed striking variation, showed little that consistently distinguished effective from ineffective writers. Information planning in assembly and whole-text review in knot-tying characterized some effective writers.

Analyses of protocols in terms of the types of problems writers defined showed that the primary distinction between effective and ineffective writers lay in how they dealt with selecting information from their base of procedural task knowledge. Effective writers tended to define more problems related to selecting information (e.g., assessing informational sufficiency, level of detail, need for feedback). They also tended to test the adequacy of information by anticipating users' responses at choice points in their emerging texts. Ineffective writers either failed to consider problems of selecting information or considered them but applied predetermined evaluation criteria. Although these criteria were often user-based, they lacked the flexibility and specificity of constructing the response of imagined users to particular alternatives. Results are discussed in terms of instruction writers' ability to represent and balance conflicts between competing sets of criteria, such as being informative and being succinct.

Results of the revising phase showed that watching a user helped all writers detect informational ambiguities in their instructions.

Table of Contents

List of Tables

List of Figures

I. INTRODUCTION

The most basic and one of the most important forms of human communication is an instruction that directs action. Of particular contemporary interest are written instructions for procedural tasks, such as assembly and trouble-shooting. In the military, for example, many jobs are structured around manuals or other documents that describe the step-by-step actions needed to perform procedural tasks. Despite the importance of the text in these situations, numerous sources indicate that written procedural instructions are inefficient and ineffective for the intended user (General Accounting Office/DoD, 1979; Duffy, 1982).

This study explores the process of writing procedural instructions and, in particular, seeks to identify and describe writing processes associated with more and less effective instructions. Our ultimate goals in this study were twofold: Practically, we sought ways to improve procedural instructions by identifying effective ways to write them. A description of effective writing processes could then inform process-based writer training or aiding programs. Theoretically, we sought to contribute to the understanding of composing processes by looking at an area that has been little studied: technical writing -- in particular, writing procedural instructions. By exploring differences among writers, we hoped to identify potentially important variables for future studies of technical writing.

A process-based approach to improving procedural instructions complements the more traditional, product-based approach. The product-based approach seeks to manipulate experimenter-controlled text features that are hypothesized to affect the processes of instruction users (Dixon, 1982; Smith & Goodman, 1982; Stone & Glock, 1981; Reder, Charney, & Morgan, 1984). The results of these manipulations ultimately inform text design guidelines that show what an effective text should look like on a number of dimensions (e.g., Hartley, 1981; Felker et al., 1981). Process-based guidelines are a potentially important complement to product-based guidelines because processes are generative. They should transfer across a range of specific user and task variables on which effective product features depend.

Our broad theoretical framework for examining instruction writing is the models of composing processes developed by Hayes and Flower (1980) and by Collins and Gentner (1980). Viewing composing as problem-solving, these models characterize writers in terms of how they define the writing problem for themselves, in terms of the writing goals that come out of this definition, and in terms of how they define and manage constraints on composing. Writers are seen as drawing on and combining different types of knowledge to define problems, goals, and constraints: knowledge about the topic of the composition, about the intended readers, and about language and text conventions.

For procedural instructions, these types of knowledge translate into: knowledge about the procedural task, about the intended user, and about the text, as schematized in Figure 1.

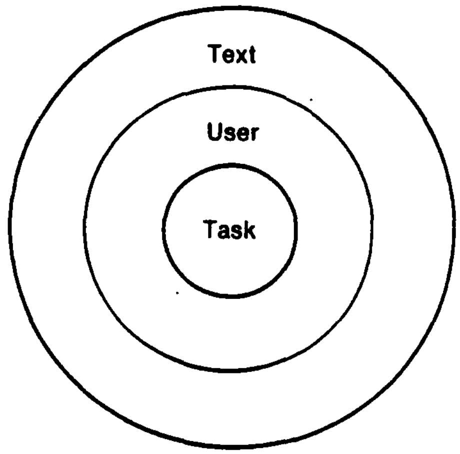


Figure 1. Types of Knowledge Required to Write Effective Procedural Instructions

Primarily using methods of protocol analysis, researchers in this framework (e.g., Flower & Hayes, 1981; Hayes, 1984; Odell & Goswami, 1982; Scardamalia, 1984; Halpin, 1984) have found that more skilled writers:

- define more reader-centered or rhetorical goals for themselves

- define more abstract, hierarchically structured representations of the writing task, which allow them to

- attend to more global levels of text in composing and revising, and to

- manage constraints systematically, as through extended planning.

Less skilled writers, on the other hand, tend to define the writing task more linearly and to follow rule-based or topic-centered goals as opposed to rhetorical goals. We expected that the same kinds of differences might emerge between more and less effective writers of procedural instructions.

To determine what processes are effective, the behaviors involved in composing and revising procedural instructions must be related to the use of those instructions. Thus, our immediate aims were:

1. to identify more or less effective writing processes in terms of how well the instructions work for users;

2. to link effective writing processes to specific features of the instructions produced;

3. to determine how feedback from the instruction users affects the nature and effectiveness of the writers' processes and products.

To accomplish these objectives, we conducted a study that (1) observed writers writing instructions, (2) tested the instructions on users, (3) observed writers revising their instructions while watching users, and (4) tested the revised instructions on users.

After identifying more and less effective instructions based on users' performance, we looked for sources of effectiveness in (1) the instructions themselves, looking at text features hypothesized to influence effectiveness, and (2) the writers, looking at a variety of processes found pertinent to good writing in previous research. These processes were characterized in terms of (1) summary measures of writers' surface behaviors, (2) types of changes writers made, and (3) categories of comments writers made in thinking aloud.

II. METHOD

Design

Sixteen writers participated individually in two composing sessions: one writing session and one revising session. The interim between writing and revising ranged from 15 to 19 days. During the writing session, writers produced a set of instructions for one of two procedural tasks: tying a necktie in a four-in-hand knot (eight writers) and assembling a 14-piece car, given a set of 39 pieces from the Fischertechnik 100 building kit (eight writers). Writers were assigned randomly to each task. During the revising session, writers revised their own instructions, first without, then with feedback in the form of a videotape of someone using their instructions.

During the interval between writing and revising, each writer's instructions were tested on ten naive users, all of whom were videotaped. The user whose time and errors were most representative of those ten was chosen to provide feedback during the revising session.

The writers in each task group also were assigned randomly to one of two methodological conditions: thinking aloud and writing aloud. Within each task group, the four writers in the thinking-aloud subset were asked to verbalize their thoughts

while they composed. The four writers in the writing-aloud group were asked to say aloud what they were writing down while they composed. The assignment of writers to thinking-aloud and writing-aloud conditions was maintained over both the writing and the revising sessions.

The thinking-aloud condition allows us to collect protocols providing a partial record of what writers are attending to while they compose instructions. The writing-aloud condition served first as a control, enabling a comparison between writers who did and did not think aloud. Such a methodological control is frequently recommended in the use of verbalized thought protocols, because it helps to detect any gross distortions of subjects' composing processes attributable to the added task of thinking aloud. The writing-aloud condition also enabled us to record the sequence in which the writing-aloud subjects produced and made changes to their text, just as we could with the thinking-aloud subjects.

The Procedural Tasks

Tie-tying and item assembly exemplify two types of procedures: a procedure with continuous actions and a procedure with discrete actions. A continuous procedure is one in which the boundaries between the action segments are difficult to determine. It involves high levels of motor skill and automaticity -- thus, the expert performs the actions in a more coordinated and highly chunked manner than the novice. A

discrete procedure involves low levels of motor skill and little automaticity -- the nature of the actions, the divisions between them, and the degree to which they are separable all change relatively little from the novice to the expert performer.

We chose the tie-tying task to exemplify a continuous procedure for several reasons: It is simple to manage in the laboratory and can be reliably scored. Writer subjects who are already task experts are easy to find. User subjects who are novices are also easy to find.

We chose the Fischertechnik task to exemplify a discrete procedure because tasks from that kit have been used in previous research on procedures and are demonstrably feasible and describable (Stone & Glock, 1980; Bieger & Glock, 1982; Crandell & Glock, 1981; Baggett & Ehrenfeucht, 1982).

We reasoned that these two types of tasks would place different sets of demands on instruction writers. We wanted to see how writers responded to these demands and also whether effective instruction writing processes were task-dependent.

A descriptive scheme was developed to characterize each procedural task, for use in assessing the accuracy of writers' instructions and in scoring the performance of instruction users. The task description for tie-tying, shown in Table 1, consists of an ordered series of steps. Each step is composed of (1) an operation -- e.g., moving the wide end of the tie with reference

TABLE 1. Task Description for Tie-tying Procedure

OPERATION	STATE
1. Puts tie around neck.	Ends hang in front of chest. Tie close to back of neck. Faces out.
2. Pulls ends.	Wide end hangs twice length narrow end.
3. Crosses wide end over narrow end.	Faces out.
4. Crosses wide end under narrow end.	Lower face wide end in.
5. Crosses wide end over narrow end.	Lower face wide end out.
6. Moves wide end under and up behind loops of tie. Holds loops with subordinate hand, at mid-chest.	Lower face wide end in.
7. Moves wide end out over loops. Holds loops with subordinate hand.	Lower face wide end up.
8. Moves wide end down through outer loop of knot. Holds knot with subordinate hand, at mid-chest.	Lower face wide end out.
9. Pulls wide end down all the way. Holds knot.	Knot tightens.
10. Moves knot up. Holds narrow end.	Knot at collar.
11. Adjusts tie.	Knot tight. Knot flush with collar. Ends parallel. Ends flat. Narrow end under wide end. Narrow end shorter than wide end by a few inches.

to the narrow end -- and (2) a state, describing the result of an operation.

The task description for the car assembly, shown in Table 2, consists of (1) an unordered list of steps -- characterized as piece-selection operations (e.g., "select a wheel") and piece-connection operations (e.g., "connect the wheel to the axle"); and (2) a graph of the ordering dependencies between the various connection operations. An exploded diagram of the assembly appears in Figure 2, with an inset picture of the finished car. Distractor pieces, in the set of 39 from which the car was constructed, were chosen to be maximally similar to the 14 target pieces.

Writer Subjects

Sixteen adults holding professional jobs served as instruction writers. They were paid $10.00 for their participation. These subjects were sampled from a population we termed "generally experienced writers": those who, while not necessarily professional writers, do write as some part of their jobs and produce different kinds of texts. We considered this the best population to provide baseline data on instruction writing. It clearly excludes basic, beginning, and novice writers, who are noted for having difficulties at various levels of composing. We were not interested in these kinds of writers.

TABLE 2. Task Description for Assembly Procedure
(Constructing a Model Car)

Unordered List of Selection Operations and Connection Operations for Model Car Pieces:

SELECTION OPERATIONS

ST - steering wheel top
SM - steering wheel middle
SB - steering wheel bottom
C - clip
R - rod (axle)
L - link beam (or front axle)
P - base plate
H - tail piece hook
A1 - back (stub) axle
A2 - back (stub) axle
W - wheel (1, 2, 3, 4)

CONNECTION OPERATIONS

L-W1 - (link beam to front wheel)
L-W2
L-R
R-P
R-SB
SB-SM
R-ST
SM-ST
A1-W3
A2-W4
A1-P
A2-P
H-P

Ordering Dependencies Among Connection Operations:

← ordered →

unordered

Piece Connections

L-W1
L-W2
L-R → R-P → R-C → R-SB → [SB-SM → R-ST]
[SB-SM
SM-ST]
A1-W3
A2-W4
A1-P
A2-P
H-P

(Note: The two sets of brackets indicate two optional orderings of the connections within the brackets. The steering wheel can be assembled piece by piece on the axle rod (R) or can be assembled independently and then placed on the rod.)

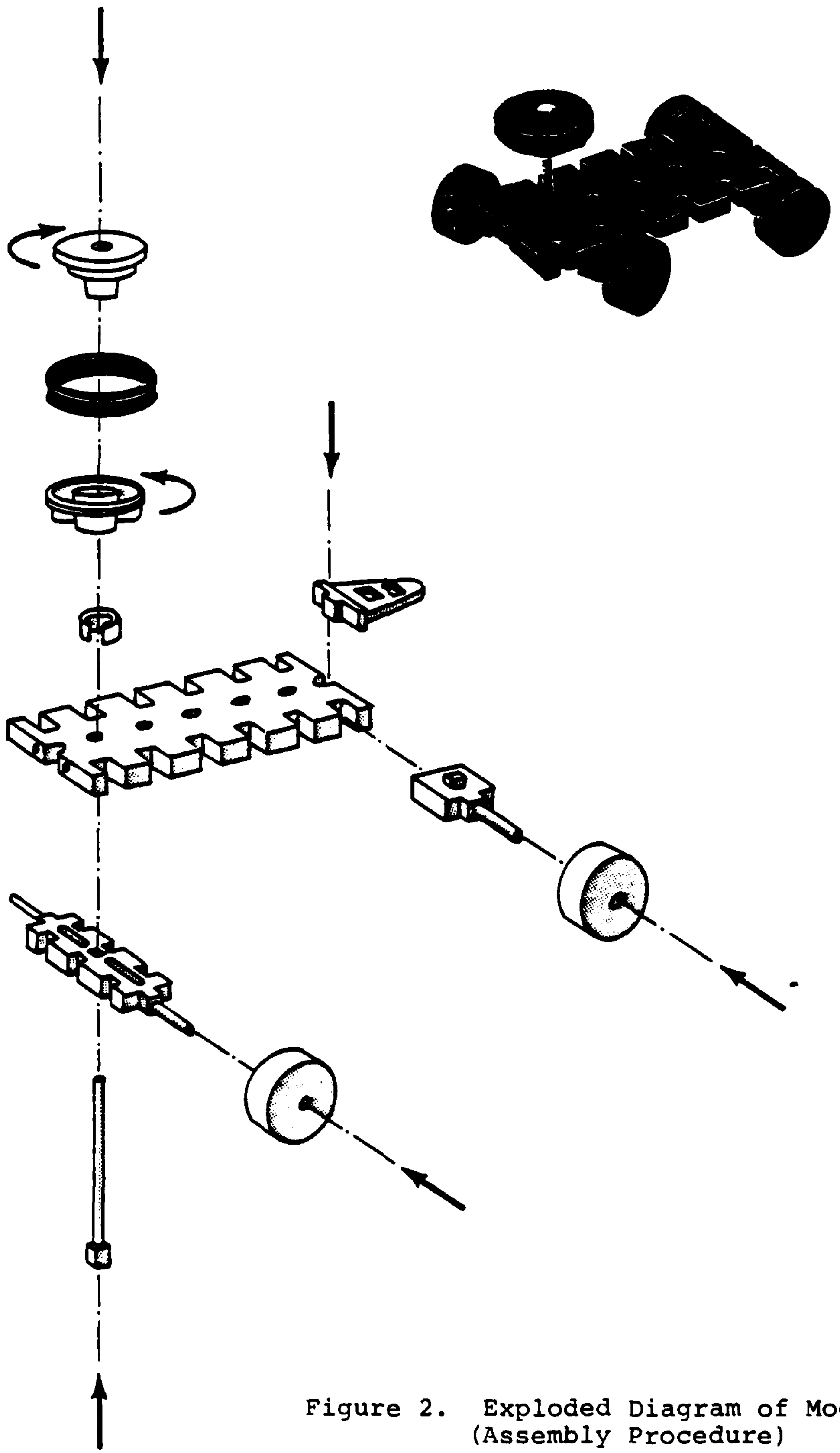


Figure 2. Exploded Diagram of Model Car
(Assembly Procedure)

To identify "generally experienced writers," we modified criteria developed by Peterson (1983).

(1) Writers must have at least a bachelor's degree plus two years of work experience in a professional job that requires some amount of writing.

(2) They must write regularly -- defined as five hours per week of writing and editing on the average.

(3) They must write different kinds of texts -- specified as at least four types, with types distinguished as reports, memos, business letters, proposals, brochures, forms, stories, press releases, etc.

In addition, writers selected for the tie-tying task were all a priori task experts: They said they knew how to tie a four-in-hand knot and were able to do so correctly and without hesitation. Writers selected for the assembly task had not had prior experience with the Fischertechnik car assembly.

User Subjects

Each set of original instructions produced by the 16 writers was tested on ten naive users. Each set of revised instructions was tested on from six to eight naive users. To obtain naive users, we selected subjects who indicated they had no experience using the Fishertechnik kit or tying ties.

User subjects were drawn primarily from two local universities and from surrounding offices. To ensure a minimum reading level, we selected subjects who were either enrolled in a university or had a college degree. Subjects were paid $5.00 for their participation.

Procedure for Observing Writer Subjects

General directions. Writers each spent one session writing and one session revising. The task environment was relatively unconstrained. The basic directions were to write instructions for someone who had "never done the task before -- an adult with normal reading ability." Writers were told that their instructions would actually be tested out on this type of person, and that they would have a chance to see a videotape of the results. Writers were told they could take as long as they wanted and could approach the writing task however they liked.

Writers for the tie-tying task had a tie available while writing. Writers for the assembly task had both the set of 39 pieces and the exploded diagram shown in Figure 2. Writers for this task were first trained to mastery in the procedure by disassembling and assembling the car with the aid of the exploded diagram (this diagram avoids biasing writers toward any particular order of construction).

All sessions were audiotaped, and all written products were collected at the end of each session, including notes, outlines, and successive drafts.

Thinking aloud and writing aloud. Writers in the thinking-aloud condition were given the thinking-aloud directions described by Peterson (1983), based on those of Hayes and Flower (1980). These directions asked writers to try to verbalize everything they were thinking about while they were writing. As practice in thinking aloud, writers wrote a short set of directions for finding a given room in a building. During practice, writers were given feedback to encourage fluency in verbalizing their thoughts.

Writers in the writing-aloud condition were asked to say out loud what they were writing as they wrote. These writers were given the same practice with feedback as in the thinking-aloud condition.

Revising. Writers were asked to revise their instructions before they watched a videotape of a representative user and then again while they watched a videotape. They were asked to make whatever changes, if any, they felt were needed "to make your instructions more effective." As with the writing session, no time limits were placed on revising. Writers were free to revise and to use the videotape in whatever manner they chose. Writers switched pen color when revising with the videotape so that we could easily distinguish changes made before and after feedback. In addition, a second copy of each writer's instructions was available if he or she wanted to use it.

User Testing Procedure

Each user read aloud and followed one set of assembly and one set of tie-tying instructions (counterbalancing order of presentation). Users were asked to follow the instructions in whatever way felt comfortable to them -- they could pause, go back, reread, etc. In addition, they were asked to read aloud and to verbalize any difficulties they might run into. They were instructed to touch a bell when they had reached the end of the task. Users were videotaped as they performed.

III. RESULTS

Did Instructions Differ in Effectiveness?

Looking at Users' Performance

Scoring

Each user was scored on performance time and accuracy. Performance time was determined by measuring from the time the user began reading instructions until the time the user touched a bell, signalling the end of a trial. Accuracy was determined by matching users' performance step-by-step with the relevant task description (see Method).

The accuracy score for tying a tie indicated whether the user had achieved a four-in-hand knot. Users received a score of "2" for a topologically correct knot, "1" for an incorrect knot, and "0" if they failed to tie a knot (or if the knot fell apart or the tie was looped but not knotted).

There were two accuracy scores for assembling a car: (1) the functionality of the final assembly and (2) the number of discrete structural errors. Functionality, a measure employed by Baggett (1983), reflects the overall success of the assembly. Users received a score of "2" for a fully functional car, "1" if either the steering wheel worked or the car wheels worked but not both, and "0" if neither the steering wheel nor the wheels

worked. Number of errors, a measure used by Stone and Glock (1981) and elaborated by Schorr and Glock (1983), reflect the discrete errors of various types that occur in the final assembly (orientation of parts, location of parts, missing connection, wrong parts used, adjustment of parts). An assembled car could have errors (e.g., leaving off the trailer attachment) and still be fully functional on the criteria we defined. Therefore, even though the two measures were correlated significantly ($\underline{r} = -.80$, $p < .01$), we examined performance in terms of both.

Time and accuracy measures for tying a tie were based on the users' first-trial performance. (While users sometimes backtracked on the tie, they always did this early in the task. Backtracking was difficult after the knot-forming movements were reached. To correct errors after this point, users had to start the instructions over, thus beginning a new trial.) Assembly users performed only one trial -- they always backtracked if they ran into problems before they finished the instructions.

Findings

Assembly. Table 3 shows the mean scores for functionality, error, and time for the ten users of each of the eight sets of assembly instructions.* As shown in Table 3, different instructions tended to produce different mean levels of functionality -- from a low of .80 to a high of 1.9

---

*Eleven users performed the assembly task with Writer 3's instructions.

TABLE 3. User Performance Data for Assembly Instructions

	PERFORMANCE MEASURE			
ASSEMBLY INSTRUCTIONS WRITER #	Mean Functionality	Mean Total Errors	Mean Time (Minutes)	(% Users with Fully Functional Assembly)
15	1.9	2.0	13.28	(90)
16	1.5	3.1	16.87	(60)
10	1.4	3.0	22.27	(60)
6	1.4	3.8	17.11	(60)
2	1.3	4.2	18.71	(60)
7	1.3	5.0	21.57	(50)
9	.9	6.5	20.67	(30)
3	.8	6.5	16.20	(30)

(F[7,73]=1.95, p=.075). Instructions clearly varied in terms of the mean number of total errors, producing on the average from 2.0 to 6.5 errors (F[7,73]=3.31, p=.004).

These two types of data supported our expectations that writers' instructions would impact on accuracy measures of user performance. Writers were ranked almost identically on both accuracy measures, with Writer 15's instructions clearly producing the most accurate performance and the instructions of Writers 9 and 3 leading to the least accurate assemblies.

The different sets of instructions also tended to differ in terms of mean performance time (F[7,73]=1.85, p=.089). It should be noted, however, that the performance times of users within any particular group showed considerable variation. This was especially true for the <u>less</u> accurate instructions, where some users continually reread and retried steps they did not understand while others gave up on those steps.

The ranking of writers according to mean performance time is generally consistent with the accuracy rankings (that is, instructions that produced more accurate performance also took less time to execute). There were two exceptions: Writer 10's instructions produced relatively accurate performance but took the longest time; Writer 3's instructions, although least accurate, took the second least amount of time.

Using arbitrary cut-offs on the distributions of mean functionality, error and time data, we defined three levels of effectiveness for the writers of assembly instructions as follows:

Most efffective - Writer 15

Moderately effective - Writers 16, 10, 6, 2, 7

Least effective - Writers 9, 3.

The resultant groupings were more heavily weighted in terms of the accuracy than the time measures. The reliance on accuracy to define effectiveness of instructions seemed the most reasonable course given the ambiguous meaning of performance time data, as mentioned above. The relative effectiveness of the various sets of instructions is also indicated by the data in Table 3 that show the proportion of users in each group who achieved a fully functional assembly. Nine out of the ten users who followed Writer 15's instructions produced a fully functional assembly. Only six of the 21 persons who followed the instructions of Writers 3 and 9 were successful.

The specification of different levels of effectiveness provides a point of departure for subsequent discussions of factors potentially associated with relatively more or less effective instructions. Consequently, in later sections when we examine the instructions per se and the processes writers use when generating instructions, we will focus discussion on Writers 3, 9, and 15. The first two of these stand out most clearly as ineffective, while the third clearly was the most effective.

Tie-tying. Table 4 shows the mean accuracy and time scores for the ten users of each of the eight sets of tie-tying instructions. Different instructions produced different mean levels of accuracy ($F[7,72]=3.04$, $p=.007$), with Writers 14, 4, and 12 producing lower means than the other writers.

If we look just at the proportion of users who got a topologically correct knot (by collapsing the two levels of incorrect performance) the difference among writers becomes even clearer: 15 of the 20 persons who followed the instructions of Writers 11 and 13 correctly tied a four-in-hand knot; only nine of the 30 persons who followed the instructions of Writers 14, 4, and 12 did so.

Instructions also clearly varied on mean performance time ($F[7,72]=17.54$, $p <.0001$). Writers 13 and 8 produced the fastest mean times among users while Writers 11, 5 and 12 produced the slowest times.

It is interesting to note that two sets of instructions that yielded relatively long performance times also produced relatively more accurate performance (Writers 11 and 5). Conversely, two sets of instructions that yielded performance times in the medium range produced relatively less accurate performance (Writers 14 and 4). There appear to be two principal reasons for this result.

TABLE 4. User Performance Data for Tie-tying Instructions

	PERFORMANCE MEASURE		
TIE-TYING INSTRUCTIONS WRITER #	Mean Accuracy (Correctness of Knot)	Mean Time (Minutes)	Users with Fully Correct Knot (%)
11	1.7	6.7	(80)
13	1.6	3.0	(70)
5	1.6	7.2	(60)
8	1.5	2.7	(60)
1	1.5	4.0	(60)
14	1.0	4.3	(40)
4	1.0	3.8	(30)
12	.5	7.2	(20)

For the relatively more accurate instructions, performance time covaried with length of instructions. Thus, Writers 11 and 5, who were generally effective, composed longer instructions than other writers (over 700 words, compared with between 230 and 500 words for other writers). Much of the extra length consisted of introductory and motivational statements, a type of text that was lacking in other writers' instructions and that seemed to arise from these two writers' broader conception of their task. These two sets of instructions took longer to read aloud, and this extra reading time contributed to increased performance time.

For the relatively less accurate instructions, users frequently made a critical error -- e.g., one that caused the tie's knot to fall apart -- before they finished reading or before they got to the final adjustment operations of the procedure. These users might then complete a trial in <u>less</u> time than users who were successful.

For these reasons, we again felt that the accuracy measures would provide more valid indications of effectiveness than the time measures. For purposes of further discussion, therefore, we focused on accuracy and grouped writers for the tie-tying procedure into three different levels of effectiveness as follows:

Most effective - Writers 11, 13

Moderately effective - Writers 5, 8, 1

Least effective - Writers 14, 4, 12.

In this section, we have analyzed user performance data to define relatively more effective and relatively less effective instructions for both tasks. Having made these distinctions, we next examine the characteristics of the instructions per se that might account for the observed differences in user performance.*

*Writers assigned to the writing aloud condition (Writers 2, 3, 6, and 9 for assembly; Writers 11, 12, 13, and 14 for tie-tying) did not differ noticeably in the measures of effectiveness we have discussed. Nor did they differ in any immediately obvious ways on the various measures of text features or writing process that we discuss in subsequent sections. They did take slightly longer, on the average, to finish the writing session.

What Produced the Differences? Looking at the Instructions

To explain the differences between relatively more and less effective instructions, we first looked at the instructions themselves. Based on prior research on procedural instructions and on our intuitions about what makes instructions work, we hypothesized that effectiveness would depend on certain text features:

- the factual accuracy of instructions
- the degree of hierarchical structure
- sentence readability and format variables
- the sufficiency of information content.

In the following sections, we describe how we operationalized each of these features and present our findings.

Factual Accuracy

The first obvious feature of instructions that might account for the degree of relative effectiveness is factual accuracy. Did instructions differ in content errors? For example, were steps omitted? Did an instruction tell the user something wrong, e.g., to make three loops with the tie or to attach the wheels to the wrong location on the car?

<u>Method</u>. To assess the accuracy of instructions, we mapped each set of instructions against the steps in the relevant task description -- tie-tying or assembly (see Method). We scored instructions for inaccuracies of the following types:

(1) omissions of a step represented in the task description, (2) intrusions of a step not represented in the task description, and (3) inclusion of wrong information about a step. In addition, for the assembly task we looked for violations of the ordering dependencies between operations.

Findings. Virtually no inaccuracies of these types were found. Thus the errors seen in the users' performance were not due to superficially inaccurate instructions at this gross level. Of course, this level of analysis assumes a certain degree of omniscience; when we scored instructions, we knew what to look for. We could not tell from this relatively superficial examination whether users would recognize each step or would interpret the information correctly.

Degree of Hierarchical Structure

Several researchers (Meyer, 1975; Kieras, 1979; Kintsch & van Dijk, 1978) have reported that a discriminating feature of effective text is its organization. For procedural text, Smith and Goodman (1982) showed that hierarchical organization was easier for text users than was linear organization. Thus, the next hypothesis we examined was whether differences in the hierarchical structure of instructions were related to effectiveness.

Method. The degree of goal hierarchy in each set of instructions was determined semantically, by applying a text grammar for procedural instructions developed by Gordon, Munro,

Rigney, and Lutz (1978). We modified this grammar to enlarge and clarify the definition of "goal," a major constituent of the grammar. We borrowed a definition of "goal" from Smith and Goodman's (1982) account of hierarchical elements in procedural instructions. In Smith and Goodman's account, hierarchical elements include functional explanations (e.g., "This c-shaped clip holds the steering rod against the chassis.") and structural explanations (e.g., "There are three components in the steering wheel assembly.").

The degree of higher-level goals was, to some extent, reflected in the graphical formatting of the instructions--more hierarchical instructions tended to have more headings at more levels. However, even instructions formatted as continuous prose paragraphs, or as a list of numbered steps, varied in hierarchy depending on the semantic nature of the statements introducing paragraphs or steps.

Findings. The instructions for the tie-tying procedure were all fairly linear. Aside from naming the procedure in a title or introductory statement (e.g., "How to tie a tie"), almost none of the instructions contained goals or explanations beyond the basic-level, step-by-step task information.

There were two exceptions: The instructions produced by Writer 5 contained goal statements which appeared in headings and which grouped sets of steps, for example, "Making the X," "Making the Knot." In addition, Writer 11 produced instructions with low-level hierarchical statements at two points and a global

statement at one: These statements were not goals but rather structural explanations of the cyclical pattern that relates one task operation to another (e.g., "This next step reverses the previous step."). Neither type of hierarchical feature bore a decisive relation to the success of the instructions. While Writer 5 and 11's instructions were effective, other effective instructions did not contain hierarchical elements.

The assembly instructions were more hierarchical than the tie instructions. All assembly instructions contained some local goal statements. These statements precede single connection operations and pairs of single connection operations: "To attach the tail piece, snap the peg on this piece into the hole at the back end of the base."; "To attach the back axles slide the knob of each axle into the back groove on each side of the chassis." Many assembly instructions contained higher-level goal statements. These statements grouped sets of three or more connection operations: "To form the steering wheel, you will connect four pieces..."; "To assemble the front end of the vehicle, do the following steps."

Table 5 shows the degree of goal hierarchy found in the eight assembly instructions. The counts shown include all goal statements except the highest task goal (to assemble a car), which was present in every set of instructions.

Looking at the total number of goal statements, we find that the instructions varied widely -- from three to 16 explicit statements. However, this dimension did not consistently

TABLE 5. Number of Goal Statements at Different Hierarchical Levels for Each Set of Assembly Instructions

	Level of Statement				
ASSEMBLY INSTRUCTIONS WRITER #	Grouping one connection operation	Grouping two connection operations	Grouping three connection operations	Grouping four or more connection operations	Total Statements
2	4	2	2	1	9
3	1	2	0	3	6
6	6	3	3	4	16
7	3	2	1	0	6
9	1	1	1	0	3
10	2	3	2	1	8
15	3	3	0	1	7
16	2	2	0	1	5

distinguish more from less effective instructions. Although the ineffective instructions of Writer 9 had the fewest goal statements, those of ineffective Writer 3 were in the middle of the range, as were those of the most effective writer, Writer 15. The best example of effective instructions with relatively little hierarchy is the set produced by Writer 16, which had the second fewest goal statements. As is also evident from Table 5, the statements in ineffective Writer 3's instructions tended to be high-level, while those in effective Writer 15's tended to be low-level. (The instructions produced by Writers 16, 3, and 15 are shown in Appendix A.)

These findings cast doubt on the hypothesis that more hierarchical instructions would be more effective. This hypothesis seemed reasonable because goals and explanations help constrain the instruction user's choice of actions. However, the impact of goals on users' performance appears to depend on the kind of information that the goal provides. Local goals typically provide specific constraints on connection operations. For example, a functional explanation like "the steering wheel has two pieces that screw together" constrains the selection of pieces much more than the independent piece descriptions some writers used, such as "...the piece with the large hole and grooves inside." Users who got instructions with the functional explanation were better able to pick out the two steering wheel pieces than users who got instructions with independent descriptions. But more global goals, such as "The next part of the task will be putting together the front end of the car: the

steering rod, the steering wheel, and the front wheels and axle," typically don't add more in the way of operational constraints.

Global goals do serve to group the connection operations into convenient chunks for readers, but this convenience either had little effect on the immediate performance measures we used, or it was overridden by concrete informational deficits at the basic step level.

Sentence Readability and Format Variables

Prior research implicates a variety of sentence and format features as influential in the effectiveness of written language: surface variables such as text length and sentence length; syntactic variables such as active vs. passive; order-of-sentence-mention variables; and graphical format variables, involving paragraphing and headings (see review in Jonassen, 1982; Holland, 1981). Technical writing guidelines typically recommend the use of short active sentences and frequent headings and paragraph divisions to make reading easier (Felker et al., 1981; Kern, Sticht, Welty, & Hauke, 1976; Houp & Pearsall, 1983). The next question of interest, then, was whether effective and ineffective instructions were distinguishable on any of these features.

Method. We examined the most salient and easily measured of these features: text length (in number of words), mean sentence length (in number of words), number of paragraph or graphical divisions, and mean number of sentences per division. The

results are summarized in Table 6. It is clear that both effective and ineffective instructions display a range of values on each of these measures. No measure perfectly distinguishes effective from ineffective instructions.

However, when we look at each task separately, we find that ineffective instructions do tend to be associated with particular features. In the tie-tying task, the three least effective instructions (Writers 4, 12, and 14) had fewer graphical format divisions than did the more effective instructions. While effective instructions varied in whether they used paragraphs or numbered steps, they all tended to have shorter steps and paragraphs than the three ineffective instructions. Our impression on viewing users' performance is that long paragraphs per se did not impede performance. That is, users did not tend to lose their place while reading a paragraph or express dismay at its density. It may be that the longer paragraphs reflect the writers' underlying tendency to ignore user-related aspects of instructions at many levels.

In the assembly task, the two least effective instructions (Writers 3 and 9) tended to be shorter than the rest. They contained, respectively, the fewest and the third fewest words in the set of eight instructions. Moreover, the instructions with the second fewest words (Writer 7) ranked third from the bottom in user effectiveness. The relatively short length of the less effective instructions appears to reflect the amount of

TABLE 6. Selected Sentence and Formal Features of Instructions

INSTRUCTION WRITER # / FEATURE		Total Number of Words	Total Number of Sentences	Mean Words per Sentence	Total Number of Paragraphs or Divisions*	Mean Sentences per Paragraph
TIE						
MOST	(11)	715	40	17.9	14	2.9
	(13)	284	14	20.3	7	2.0
MODERATE	(1)	453	25	18.1	14	1.8
	(5)	702	40	17.6	9	4.4
	(8)	234	9	26.0	7	1.3
LEAST	(12)	396	22	18.9	2	10.5
	(14)	496	20	24.8	4	5.0
	(4)	357	22	16.2	3	7.3
ASSEMBLY						
MOST	(15)	663	49	13.5	40	1.2
MODERATE	(16)	789	39	20.2	18	2.2
	(6)	1,050	75	14.0	46	1.6
	(2)	1,418	77	18.4	10	7.7
	(10)	875	74	11.8	37	2.0
	(7)	452	26	17.4	2	13.0
LEAST	(9)	501	24	20.9	12	2.0
	(3)	301	34	9.0	30	1.1

* Heading counts as a division of a complete sentence. Items in list count as separate divisions if they are sentences.

information they contained, a feature we examine in the next section.

A final sentence feature that deserves mention here is syntactic form. Although we did not systematically analyze syntax features for each set of instructions, our impression, from viewing users' performance and comparing it with the text of the instructions, was that these features by themselves had little effect. For example, the tie-tying instructions that produced the most accurate performance (Writer 11) contained numerous multi-clause sentences of the following sort:

> "The skinny end which is now grasped by the right hand, should be held at a point five inches above the skinny point at the skinny end."

> "The fat end is transferred to your left hand and the skinny end is transferred to your right hand."

The language of such sentences appears indirect and difficult, featuring passive voice and multiple embedded prepositional phrases. Moreover, this writer's instructions frequently mixed indirect descriptive directions such as those above, with direct prescriptive directions -- another indication of bad style.

For the assembly task, the instructions that produced the second most accurate performance (Writer 16) were full of complex, indirect sentences with infelicitous phrases:

> The front axle is longer than it is wide and has two stems or pins with slots therein, on opposite sides, to which pins will eventually attach the wheels.

Hard-to-read structures like this were not prominent in other assembly instructions. For the performance measures we used, features of this sort appeared insignificant compared with the sufficiency of the information contained in the instructions.

Sufficiency of Information Content

Another potential difference between effective and ineffective instructions is the degree of ambiguity or gaps in the specific information given within a task step. For example, did an instruction tell users to attach a wheel to an axle, but fail to specify that the smooth surface of the wheel faces out (orientation information)? Did it tell users to attach an axle to the back of the car but fail to indicate which of four back grooves was the point of attachment (location information)? Sometimes users can infer information that has been left unspecified; if not, they make predictable errors. For example, Bieger and Glock (1982) found that the number and types of users' errors corresponded to the number and types of information omissions in procedural instructions. Thus, the next hypothesis we looked at was whether -- given that all the steps were there -- the sufficiency of information content in the instructions was related to effectiveness.

Method. We analyzed information content in terms of a sample of the categories proposed by Bieger and Glock (1982) to

characterize the critical content of procedural assembly instructions: location, orientation, action, object, description, measure, and degree. Drawn from the cases of Fillmore (1968) and the pictorial categories defined by Mandler and Parker (1976), these categories proved reliable in identifying the types of information critical to instruction users' correct performance of assembly procedures (Bieger & Glock, 1982). We elaborated this list to characterize more precisely the information requirements of the assembly and tie-tying procedures.

- For assembly, we considered location in three dimensions: Did the writer specify the location of a connection on horizontal, vertical, and depth axes? For example, "connect the axle to the back groove of the chassis" fails to indicate whether the "back groove" is the last groove on the side of the chassis -- horizontal axis -- or the single groove on one end of the chassis -- depth axis. We also considered object in terms of part and subpart. For example, "connect the axle to the chassis" does not indicate whether the point of connection is the protrusion on the side of the axle or the peg on the top of the axle.

- For tie-tying, we considered location in terms of four spatial specifiers necessary to map the

movement of the wide end of the tie in three dimensions: start-state, end-state, direction of movement, and dimension of movement (dimension is with reference to a stable object, generally the narrow end of the tie). For example, "move the wide side under" specifies the dimension of movement but fails to say where the wide side ends up (end-state). We also added the category of instrument information to characterize references to hand and finger placement (e.g., "hold the wide end with your right hand").

For each set of instructions, we identified those segments of text corresponding to the basic task steps (those given in the task descriptions for each procedure) and noted for each segment whether the information in each relevant category was present or absent. For the tie-tying instructions, we focused on the five critical task operations (3, 4, 5, 6, and 8 in the task description), since these defined the topology of the four-in-hand knot. For the assembly instructions, in addition to analyzing the selection and connection operations, we counted the number of distinct piece descriptors in selection operations. For example, "find the piece that is triangular in shape, curved at the top, and has two holes in the middle" was scored as having three descriptors, while "find the small, thimble-shaped piece" was scored as having two. Two coders reached 87% agreement on the presence or absence of categorical information in a sample of

instructions. Disagreements occurred primarily for specifications that both coders felt were ambiguous. Therefore, apparent ambiguities were always referred to a second coder.

Findings: Tie-tying instructions. The information sufficiency hypothesis was strongly supported. The three ineffective tie instructions uniformly were distinguished by omissions of certain essential spatial specifiers in the critical task operations -- start-states, end-states, direction of movement, and dimension of movement with reference to a stable object (generally the narrow end of the tie). In addition, the results pointed to an unexpected source of ineffective instructions -- excessive detail. Excessive detail meant: (1) inclusion of spatial details beyond the essential spatial specifiers or (2) inclusion of hand and finger details beyond those found in the more effective instructions.

Omitting essential spatial information makes the instructions ambiguous. Instruction users may move the tie in the wrong direction or end a move at the wrong position. These errors violate the basic topology of knot-tying, and the user ends up without a four-in-hand knot. For example, directions like this occurred in the tie-tying instructions:

> "Put the wide end through the outer loop." - Directional specifier omitted. Users may go down through the loop or up through the loop. If the latter, the incipient knot falls apart.

Even the effective instructions contained occasional omissions. But the three ineffective instructions were unique in that they lacked clear end-states for at least one of the five critical operations and clear dimension or direction specifiers for one of the last two critical operations. Such omissions were not found in the more effective tie-tying instructions.

Including extra spatial details also makes the instructions ambiguous. Users can't tell how to segment the operations nor how to identify the essential spatial specifiers. They may perform extra operations or choose the wrong specifiers as the essential ones. Again, users lose the four-in-hand knot. Two of the three ineffective instructions -- those of Writers 12 and 14 -- contained extra spatial specifiers in two or more critical operations. The result was overly segmented instructions. These extra specifiers did not occur in other writers' instructions.

To illustrate how inserting extra details creates overly segmented instructions, we can compare an ineffective set of instructions with an effective set, one that presents only the essential spatial information:

Writer 11 (Effective Instructions)	Writer 14 (Ineffective Instructions: over-segmented in Step 4)
Step 4	Step 4
The left hand will move the fat end of the tie to the right side behind the skinny end.	Bring your left hand holding the large end of the tie under the cross of the tie (to the right), passing this end up on the right hand side of the tie-cross, through the circle formed by the right hand and arm, the tie, and the chest.

The two writers' versions of Step 3 (from our conceptual task description) corresponded almost exactly in terms of spatial and operational information. But in Step 4, presented above, Writer 14 incorporated several more spatial and operational specifiers than Writer 11 (those specifiers are underlined in the examples.)

In addition, the two overly segmented sets of instructions (i.e., those of Writers 12 and 14) both contained extra specifiers for hand and finger placement, far more than those in the most detailed of the effective instructions. For example, Writer 12 instructs us to "Remove right index finger from front of crossing, place it inside of outer loop of crossing against thumb of right hand." These detailed instrument specifiers appeared to add to users' confusion. Users would reread them several times and stress during debriefing that they had found them hard to follow.

In summary, effective and ineffective tie-tying instructions differed in the type and level of information provided. Ineffective instructions omitted certain spatial specifiers, such as end-states. In addition, two of the ineffective instructions were excessively detailed with respect to other spatial specifiers and instrument information.

Findings: Assembly instructions. Like the tie-tying instructions, the two ineffective assembly instructions were distinguished by omissions of information. The chief omissions were of: (1) selected location and orientation information in the connection operations and (2) sufficient piece descriptors in

the selection operations. Users' errors were predictable in terms of these omissions. For example, when insufficient piece descriptors were given, users tended to choose the wrong piece.

Excessive detail was not a problem: over-segmentation did not occur and instrument information was virtually absent. Extra details of other kinds of information did not appear to confuse users. In fact, informational redundancy tended to produce better performance, particularly redundancy of the piece descriptors in "select" operations.

What Produced the Differences: Looking at the Writers

Our central question was whether we could distinguish effective from ineffective writers in terms of how they composed. From the results presented above, we can infer that writers had accurate knowledge of the procedural tasks they were writing about. Given that the principal text features distinguishing effective from ineffective instructions involved informational sufficiency and level of detail, what were writers doing to create this distinction? We looked at three kinds of data on how writers composed: surface measures of their behavior, changes they made as they wrote, and comments they made in thinking aloud. Prior research on writing and our intuitions about the writing of instructions point to these kinds of data as potentially important in distinguishing more and less effective writing.

Surface Measures of Writers' Behavior

The first question we examined about writers was whether certain gross dimensions of their composing behavior were related to the effectiveness of their instructions. Thus, effective writers might spend more time on the writing task than ineffective writers. They might spend a longer time reviewing their drafts or tend to produce more than one draft. Table 7 displays selected surface measures of writers' behavior:

- total time-on-task,
- time to first word generated in the first draft (a measure of time spent prewriting, or plannning before writing),
- time spent in after-draft review,
- number of written plans (outlines, notes) and where they occurred in the drafting sequence (before drafting, while drafting, after a draft),
- number of drafts,
- number of after-draft reviews (where the writer rereads, or makes changes, or both).

These measures provide a kind of style profile on each writer.

TABLE 7. Surface Measures of Writers' Behavior

INSTRUCTION WRITER #	BEHAVIORAL MEASURE: Time-on-task (minutes)	Time-to-first-word (prewriting time) (minutes)	Time-to-review (After Draft 1) (minutes)	Number Written Plans	Where Plans Occur (BD=Before Draft) (AD1=After Draft 1)	Number Drafts	Number of Reviews (after completing draft)
TIE							
(11)	59	<1	20	0	--	1	2
(13)	20	<1	8	0	--	1	1
(1)	42	2	20	0	--	1	1
(5)	71	5	21	1	BD	1	1
(8)	24	<1	14	0	--	1	2
(12)	64	<1	0	0	--	1	0
(14)	61	2	27*	0	--	1 1/2**	1*
(4)	17	<1	1	0	--	1	0
ASSEMBLY							
(15)	91	38	12	3	BD	1	1
(16)	75	<1	9	0	--	1	1
(6)	167	31	0	4	BD	1 1/2**	0
(2)	98	3	0	0	--	1	0
(10)	94	5	30	1	BD	1	2
(7)	73	1	15	0	--	2	2
(9)	59	<1	7	0	--	1	1
(3)	87	1	31*	2	AD1	2	1*

* For Writers 3 and 14, all or part of the review consisted of rewriting. Writer 3 spent all reviewing time rewriting the first draft. Writer 14 spent 25 of 27 minutes rewriting half the first draft.

**Writers 6 and 14 rewrote part of their first drafts.

We hypothesized that effective writers would:

- spend more time on the task -- this seemed an obvious expectation;

- spend more time in prewriting -- this was based on prior studies showing that more skilled writers plan more (Atlas, 1981; Flower & Hayes, 1981);

- produce longer plans, with more versions and changes, particularly global changes -- this was based on observations that more skilled writers tend to transform their knowledge and thus reorder content as they outline (Scardamalia, 1984);

- be more likely to review, and engage in more cycles of review -- this was based on prior studies of revising (Sommers, 1980; Beach, 1976; Flower, 1984).

Looking at the <u>time-on-task</u> data in Table 7, we find no support for the hypothesis that effective writers would take more time. Writers varied widely. Assembly writers spent more time on the average than did tie-tying writers, but time did not systematically distinguish more from less effective writers on either task. Thus, for example, the two most effective writers for tie-tying spent, respectively, one of the longest writing sessions (59 minutes) and one of the shortest (20 minutes). The

Pages 48 + 49 blanks

best and the worst assembly writers spent nearly equal time writing (91 and 87 minutes, respectively).

Looking at the allocation of time to prewriting and reviewing, again we find little to account for differences in effectiveness. The tie-tying writers tended to start writing immediately -- about five minutes was the longest prewriting time (Writer 5), while the assembly writers were more likely to plan first. However, within tasks, there is some support for either the time-in-prewriting or the time-in-reviewing hypothesis. Among tie-tying writers, it may be significant that two of the three least effective writers (4 and 12) spent little or no time reviewing, while other writers spent several minutes reviewing. Among assembly writers, time spent prewriting may be significant. Prewriting activities defined three basic composing patterns: breadth-first, spending a long time planning (Writers 15, 6), familiarization, with modest planning (Writers 2, 10), and depth-first, plunging into writing (Writers 16, 7, 9, 3). The breadth-first and familiarization patterns were found only among effective writers, while the depth-first pattern cut across levels of effectiveness. The most effective writer (15) spent the longest time in prewriting (38 minutes).

The possible role of planning in effective writing becomes clearer if we look at the number of written plans produced and where they occurred in the drafting sequence. Among tie-tying writers, only one writer (the one who spent five minutes prewriting) produced a plan, in outline form. This writer was

moderately effective. His outline was text-schematic -- that is, it represented the abstract constituents we would expect to find in procedural text ("introduction," "steps," "feedback," etc.). Among assembly writers, we find written plans in four cases -- the two writers who typified a breadth-first pattern each produced at least three versions of an outline,, the writer who spent five minutes prewriting produced a single outline, and one of the depth-first writers produced two versions of an outline after the first draft (to guide the second draft). This last writer alone was ineffective. Thus, prewriting plans are most clearly related to effectiveness. Further, since the most effective writer, 15, produced three versions of an outline, with over six times the number of words as any other outline, there is evidence to support the hypothesis that effective writers produce longer plans and more versions.

Looking at the final categories in Table 7 -- number of drafts and number of reviews -- we find little else that relates to writers' effectiveness.

Since plans were potentially the most relevant factor in assembly writing, we looked closer at the nature of the plans. Ineffective Writer 3's first outline used high-level headings to summarize the instructions she had just composed ("The Back Wheels. The Steering Wheel...,"etc.). Her second version repeated these headings, changing two words and breaking one heading into two. No reordering occurred. Effective Writer 15's first outline broke the task into parts and subparts: "Selecting

the pieces you'll need...[listed] the chassis, two back axles, four wheels...," etc. His succeeding version elaborated this outline into finer details, listing most of the major information categories needed for each task step. For example, his final outline listed all the piece descriptors to be used in the selection operations ("the chassis -- largest piece, rectangular, seven holes down the middle, grooves on either side"). He reordered once and for a relatively local choice -- putting wheels on early vs. putting wheels on later.

Writer 15 also contrasts with the somewhat less effective Writer 6. Her first outline was a text-schematic one ("Introduction...Goal...Steps...," etc). On trying to write, she found this type of outline insufficient and generated a task-analytic one, much like Writer 15's first outline. However, her succeeding outlines did not elaborate information but rather considered global reordering possibilities (e.g., arranging all "select" operations first vs. interspersing them with the relevant "correct" operations). These results suggest that more changes to plans may be related to effectiveness (Writer 15), but they do not support the hypothesis that these changes would involve global reordering.

Summary. For a short, linear, over-learned procedure like tying a tie, we infer that a thorough review of instructions after writing may be a necessary but not sufficient characteristic of effective writers. One ineffective writer did review at length. The fact that he was unsuccessful may be

related to the fact that he spent most of the review (25 out of 27 minutes) rewriting.

Why is after-draft review critical to tie-tying and not to assembly? We might infer that, because the tie-tying procedure is difficult to segment and lacks well-defined a priori task objects ("loops," "V," etc.), instructions for the procedure are especially prone to ambiguity and vagueness. Segmentation ambiguities are apparent only on a whole-draft review.

For a longer procedure like item assembly, with several ordering possibilities, we can make the following inferences: (1) Text-schematic outlines are insufficient to aid instruction writers. (2) Outlines that come after the first draft and serve to chunk or group the text are insufficient to aid instruction writers. (3) Task-analytic outlines appear more helpful if they are more detailed -- in particular, laying out the categories of information deemed necessary to specify task operations. (4) Such outlines can be efficiently generated from the top down (a breadth-first approach). However, (5) planning in the form of global reordering, exploring the possibilities that arise from a hierarchical task representation, does not appear to lead to more effective instructions. (6) Some writers can produce fairly effective instructions without an outline. These inferences are tentative, of course, based on small numbers of writers. But they translate into reasonable, empirically derived hypotheses for future exploration.

A further speculation is that task-analytic outlines help because they guide the instruction writer explicitly in selecting information from a base of task knowledge. For evidence about this potentially important selection process, as well as about the role of review, we look to data that can better reveal how writers defined the task of writing instructions. We next consider the changes writers made.

Writers' Changes

Our concern with deeper explanations of differences in effectiveness led us to examine the changes writers made to their emerging drafts. In general, writers make changes by evaluating text with respect to internally represented constraints. Thus, one way to reveal how the writer defines the writing task is by noting the kinds of changes the writer makes. For example, writers who make changes that are global in scope may be inferred to represent higher-level, more abstract constraints than writers who make only local changes.

Writers' change patterns potentially can distinguish levels of effectiveness. For example, prior studies have shown that better writers make more global changes than poorer writers (Sommers, 1980; Faigley & Witte, 1981; Halpin, 1984; Hayes, 1984).

Drawing on prior research, we hypothesized that:

- Effective instruction writers would make more extensive changes. This was based on the assumption that effective writers represent more constraints and use them more rigorously. Some studies (Beach, 1976) have shown that skilled writers tend to revise more.

- More of the changes made by effective instruction writers would be global in scope (spanning more than a sentence). This was based

on the studies cited above and on the notion that better writers represent more abstract, higher-level constraints.

- Effective writers would make more meaning changes as opposed to surface changes. This was based on the assumption that better writers conceive of their texts at more abstract, semantically based levels. Support comes from work by Faigley and Witte (1983), Daiute (1984), and Sommers (1980).

- Effective writers would be more likely to add and substitute than to delete text elements. This was again based on the assumption that poorer writers see changes as an exercise in surface editing; better writers see changes as as a chance to elaborate and transform content. Work by Sommers (1980) and by Scardamalia (1984) support this assumption.

Although the prior research we draw from relies mainly on expert-novice differences and on text other than instructions, we felt its findings about the characteristics of skilled writers would generalize to most types of effective writing.

Method. We classified the changes writers made both during and after drafting, using a classification scheme designed to capture the dimensions discussed above. This scheme contains

selected aspects of the schemes developed by Sommers (1980), Bridwell (1980), Faigley and Witte (1983), and Lutz (1983). Broadly described, these schemes categorize changes in terms of linguistic structure (adding, deleting, or substituting a word, phrase, or sentence), in terms of whether they preserve or modify meaning (defining meaning propositionally, as of text semantic structure), and in terms of temporal process (e.g., the number of lines the writer moves between changes). Specifically, we looked at:

1. the total number of changes occurring while drafting (WD), after drafting (AD), and overall;*

2. the density of changes -- the average rate of change or number of changes per words the writer produced;

3. the timing of within-draft changes -- whether immediate (within the sentence the writer is writing), delayed by one intervening sentence, or delayed by two or more sentences;

4. the linguistic scope of changes --- whether local (involving units of a sentence or less) or global (spanning more than a sentence);

5. the effect of changes -- whether or not they affect task information, in terms of the information categories we defined earlier (location, orientation, etc.).

---

*We considered only net changes in these counts -- e.g., recopying verbatim did not count as a change.

Task-related changes could be of two general types: (1) correcting wrong information, (2) selecting a different level of detail -- e.g., adding information deemed needed by a user or deleting information deemed inferrable by a user. _Non-task_ changes affect more superficial levels of text: style, readability, mechanics. These changes may, for example, consist of syntactic paraphase, lexical substitution, or grammatical correction; and

6. the _type_ of task-related change -- (1) the _operation_ (adding, deleting, or substituting information); the (2) _level_ (whether or not a task _hierarchical_ element is involved -- e.g., a goal).

Two coders reached 89% agreement in applying this classification scheme to the changes in a sample of writers' drafts. Because we were dealing with a small number of writers, some of whom made few changes, we had a second coder review all change classifications, focusing especially on cases about which the first coder was uncertain.

Findings: Tie-tying writers. Table 8 shows the results of classifying the tie-tying writers' changes. The most striking finding is that writers varied widely from one another. No dimension immediately distinguishes the three ineffective writers (4,12,14) from the more effective ones, nor the two most effective writers (11,13) from the rest.

For example, in terms of number of changes (Table 8, column 1), a measure of extensiveness, writers ranged from 11 to 109 total changes (Writers 4 and 5, respectively). Both effective and ineffective writers were represented at the high end (Writers 5, 11, and 14) as well as at the low end (Writers 4 and 13) of this dimension.

Looking at rate of change (Table 8, column 2), another measure of extensiveness, we find that overall rates ranged from one change per six words written (words retained in draft) to one change per 32 words written. Again, both effective and ineffective writers had high rates as well as low rates of change.

We can also look at how changes were distributed (Table 8, column 1) while drafting (WD cycle) vs. after drafting (AD cycle). Writers varied from 0 to 48 AD changes. Whether a writer made the majority of changes during or after drafting did not distinguish effective from ineffective writers. It may be significant, however, that the two writers who made no AD changes were both ineffective (Writers 4 and 12). This observation extends our previous findings about the allocation of time to review: thus, two of the three ineffective writers either did not review (Writer 12) or reviewed briefly without changes (Writer 4).

The rate of change within-draft provides a measure of the density of writers' changes while they are actively generating text (Table 8, column 2). Writers ranged from sparse changes

TABLE 8. Writers' Changes in Composing Tie-Tying Instructions (Totals over within-draft [WD] and after-draft [AD] cycles)

Dimension of Change

	1 Number	2 Rate	3 Timing of Within-Draft Changes			4 Linguistic Scope	5 Effect		6 Type of "Task" Change			
									(a) Operation (% Task Chgs.)			(b) Affect Hierarchical Level (% Task Chgs.)
Writer #	n (Overall) n (WD) n (AD)	(Change rate per words retained)	% Immediate	% Delayed (1 sentence)	% Delayed (2+ sentences)	% Global	% Non-task	% Task	+	-	Subst.	
Most Effective												
11	57 (38) (19)	1 in 13 (WD: 1 in 17)	63	13	24	7	68	32	88	6	6	7
13	12 (2) (10)	1 in 24 (WD: 1 in 125)	0	0	100	0	67	33	100	0	0	0
Moderately Effective												
1	52 (22) (30)	1 in 12 (WD: 1 in 15)	90	5	5	0	45	55	81	15	4	0
5	109 (61) (48)	1 in 6 (WD: 1 in 10)	87	5	8	2	62	38	83	7	10	7
8	31 (9) (22)	1 in 8 (WD: 1 in 24)	78	11	11	0	62	38	58	17	25	3
Least Effective												
12	35	1 in 12	88	6	6	0	57	43	73	27	0	3
14	88 (41) (47)	1 in 6 (WD: 1 in 11)	95	5	0	6	59	41	53	28	19	2
4	11	1 in 32	100	0	0	0	82	18	50	50	0	0

*Rewrote draft

while writing the first draft (Writer 13: one change per 125 words) to dense changes (Writer 5: one change per ten words). Writers with sparse changes have been characterized as "sprinters," and those with dense changes as "plodders" (Peterson, 1983; Hairston, pers. comm.). Prototypical sprinters, like Writer 13, save most of their revisions until after they have generated a draft -- a breadth-first strategy. Prototypical plodders, like Writers 5 and 14, revise iteratively and recursively in small chunks as they generate a draft -- a depth-first strategy. While this dimension did not distinguish effective from ineffective writers, it is notable that the only pure sprinter in the group (Writer 13) was also one of the two most effective writers for this task.

In the timing of WD changes (Table 8, column 3), the predominant pattern consists of far more immediate, within-sentence changes than delayed changes. The smallest proportions of delayed changes were found in ineffective Writers 4 and 14.

The linguistic scope of each writer's changes (Table 8, column 4) was predominantly local. A few global changes were made by three writers (no more than 7% of total changes in each case). Two of the three writers who made global changes were effective (Writers 5 and 11), while one was ineffective (Writer 14).

Considering the effect of changes (Table 8, column 5), from 18% to 55% of writers' total changes modified task information as opposed to stylistic language features. Writers toward the low end of the range were both effective (Writer 13 with 33% task-related changes) and ineffective (Writer 4 with 18%), as were writers toward the high end (ineffective Writer 12 with 43%; effective Writer 1 with 55%). Writer 1 was the only person who made proportionately more task-related than stylistic language changes.

For task-related changes, the predominant change operation for all but one writer was addition of task content (Table 8, column 6a). Writers' deletions ranged from a low of 0% of their task-related changes (Writer 13) to a high of 50% (Writer 4 who only had two task changes). The three writers with the highest overall rates of deletion were also the ineffective writers. We also looked at the function of task-related changes, and found that relatively few served to correct wrong information.

Finally, only a small percentage of the task-related changes involved hierarchical levels of task information (Table 8, column 6b). No distinction was apparent between effective and ineffective writers. Effective Writers 5 and 11 and ineffective Writer 14 made at least two such changes.

Findings: Assembly writers. Table 9 presents the results of classifying assembly writers' changes. As with tie-tying, no dimension immediately distinguishes more and less effective writers.

TABLE 9. Writers' Changes in Composing Assembly Instructions (Totals over within-draft [WD] and after-draft [AD] cycles)

Dimension of Change

	1 Number	2 Rate	3 Timing of Within-Draft Changes			4 Linguistic Scope	5 Effect		6 Type of "Task" Change			
				% Delayed					(a) Operation (% Task Chgs.)			(b) Affect Hierarchical Level
Writer #	n (Overall) n (WD) n (AD)	(Change rate per words retained)	% Immediate	(1 sentence)	(2+ sentences)	% Global	% Non-task	% Task	+	-	Subst.	(% Task Chgs.)
Most Effective												
15	38 (29) (9)	1 in 17	76	15	9	3	43	57	50	35	15	15
Moderately Effective												
16	45 (41) (4)	1 in 18	74	11	15	0	56	44	86	14	0	2
10	162 (84) (78)	1 in 5	54	17	29	6	62	38	62	20	18	10
6	174	1 in 6	43	13	44	6	75	25	72	21	9	16
2	69	1 in 20	67	24	9	0	57	43	90	5	5	3
7	157 (23) (134)*	1 in 3	91	5	4	0	46	54	56	38	6	8
Least Effective												
9	54 (49) (5)	1 in 9	60	27	13	2	55	45	65	13	22	17
3	131 (59) (72)*	1 in 3	66	7	27	3	63	37	70	22	8	12

*Rewrote draft

In terms of the number of changes, and the rates of change (Table 9, columns 1 and 2), writers varied from frequencies of 38 to 174 and had rates that varied from one change per 20 words to one change per three words. Thus, writers revised more extensively for assembly than for tie-tying, but effective and ineffective assembly writers could not be distinguished on either measure of extensiveness. The high overall rates of change in some assembly writers were attributable to the writers' rewriting their first drafts (e.g., Writer 3) -- rewriting in general led to more changes than simple revising. In terms of the rate of within-draft changes, no writer displayed the sprinter style that we saw for the tie-tying task (i.e., rates of change lower than one per 100 words).

Looking at the timing of within-draft changes (Table 9, column 3), we find that, unlike the tie-tying task, all writers made some delayed changes; a few writers made many of them. For example, 57% of Writer 6's changes were delayed and, of these, 78% entailed going back two sentences or more. The reason for more delayed changes in assembly may be that the procedure is more complex than tie-tying. It has more task operations, more information categories per operation, and more ordering options. Dealing with these complexities often requires returning to earlier points in the text and manipulating chunks larger than a sentence.

In the linguistic scope of changes (Table 9, column 4), more writers made global changes in writing assembly instructions

(five out of eight) than in writing tie-tying instructions (three out of eight). In each instance, however, the percentage of global changes was very small as in the tie-tying task. The writers who made more global changes (Writers 10 and 6) dealt with (1) higher-level formatting, such as distribution of headings, and (2) the hierarchical ordering possibilities of the task. Neither the global change dimension nor the delayed change dimension distinguished effective from ineffective writers. As with tie-tying, the tendency to change global levels appeared to reflect a difference in style of composing, not in effectiveness.

The effect of changes (Table 9, column 5) ranged from 25% to 57% task-related, a distribution very similar to that found among writers in the tie-tying task. Among both sets of writers, the tendency was to make changes that impacted more on text or style than on task information. others. This dimension was not consistently associated with effectiveness.

For task-related changes (Table 9, column 6a), addition was the predominant change operation for all writers, as was the case in the tie-tying group. Ineffective writers did not tend to make more deletions.

Task-related changes affecting hierarchical levels of information (Table 9, column 6b) were more prevalent in assembly writers than in tie-tying writers. This finding mirrors our earlier observations that the assembly instructions had more hierarchical structure. More and less effective writers were not distinguished on this dimension.

Summary and discussion. All writers changed their drafts while writing. All writers made purely formal as well as informational changes. In fact, purely formal changes comprised close to or more than half the total changes for each writer. Global changes were rare - no more than 7% for any writer. The dimensions we looked at (Tables 8 and 9) could not unambiguously distinguish effective and ineffective writers.

However, in the tie-tying task, we did see a partial distinction -- reflecting the pattern we saw in surface measures: Two of the three ineffective writers (4; 12) made no after-draft changes. Do their change patterns reveal anything about underlying approach?

One of the writers (4), who produced very sparse instructions, did in fact review but did not change or add information, despite numerous omissions of essential spatial specifiers on his first draft. We might infer that this writer defined looser criteria for information sufficiency than did other writers. He had time and occasion to correct the omissions in his text but apparently either failed to detect them or did not see them as impediments to users.

Writer 12, who produced highly over-segmented instructions, may have defined the writing problem as one of describing task information accurately and in as much depth as possible. Thus, once he had generated a draft, this problem was solved. In both cases, the writers appeared to ignore the communications aspects of instruction writing.

On the other change dimensions that we examined in Tables 8 and 9, the data did not point to any fundamental differences in problem definition between effective and ineffective writers. The change data thus fail to confirm many of our primary hypotheses: First, effective writers were not distinguished by more global changes. This is true whether we consider a linguistic definition of global changes (beyond the sentence level) or a content definition (relating to hierarchical task information). Similarly, effective writers were not distinguished by more extensive changes or by more meaning-based changes (changes to task information). There was a tendency toward fewer content deletions among effective tie-tying writers, but the differences were not clear-cut.

These findings appear to be at odds with the results cited earlier of prior research on writers' changes. The discrepancy could lie in the nature of instruction writing. Instruction writing involves fixed content -- in this case, a procedural task -- while essay writing involves open-ended content. Since open-ended content requires more creation and organization on the part of the writer, it is more likely to foster global text changes and changes that affect meaning.

Another reason for the discrepancy could be the nature of the writers in this study -- they were all experienced. Although experienced writers may vary greatly in effectiveness, they are less likely to show the dramatic process differences found between novices and experts or between skilled and unskilled

writers. Many of the differences we found in experienced writers' change patterns -- e.g., relatively more focus on stylistic changes -- were differences in personal style and strategy. They mirror the differences we found in the instructions themselves -- some were more densely hierarchical than others, some had shorter, easier-to-read sentences. These features were not related to effectiveness.

To see whether differences in problem definition and approach might distinguish more and less effective writers, we undertook a deeper examination of writers' processes through protocol analysis of the thinking aloud data.

Writers' Comments While Thinking Aloud

The thinking-aloud protocols give clues to how writers define the problem of writing instructions and thus complement the change data: There may be some changes writers fail to comment on, which nevertheless point to goals and constraints; and there may be goals and constraints the writer considers that are never manifested in changes. To interpret the protocol data, we looked at writers' comments together with the text segments they concurrently produced and the changes they made.

We analyzed the protocols as a sequence of problems posed by the writer. What problems did more and less effective writers posit to guide composing and reviewing? How did they solve these problems? What criteria governed solutions?

Based on what prior research has shown about how better writers compose (Flower & Hayes, 1980, 1981; Scardamalia, 1984; Flower, 1984), we predicted that effective writers would

- o define problems at a more abstract, hierarchical level than ineffective writers;
- o consider more planning problems to manage the writing process;
- o tend to pose more rhetorically based problems;
- o use more rhetorically based criteria.

Method. Our analysis of the protocol data followed the problem-solving model and method of protocol analysis described by Voss, Greene, Post, and Penner (1983). We divided each protocol from the eight writers in the think-aloud condition into problem segments. These segments generally contained more than one comment (defined as a main clause plus modifiers), and included in sequence a *problem* posed, *solution(s)* planned or implemented, and (optionally) an *evaluation* of the solution on one or more *criteria*. Writers did not always articulate evaluation criteria, but when they did, these criteria reflected the writers' *goals* underlying the problem segment. Writers sometimes posed problems that they never dealt with. In those cases, we considered the problems in our analysis and noted the absence of solution attempts.

The following portion of the protocol from Writer 3 illustrates how problem segments were delimited:

> //First thing I have to do is figure out what parts the person will need. So I have to figure out which parts are -- part of this, which means I really need to disassemble -- part of this. [Disassembles] //There's a lot of little parts to -- OK, what do we have to say to distinguish these parts from the other ones? OK, let's start with the red parts.// Um -- *to assemble the* -- [P] -- *assemble the* Fischer -- I don't know what to call it -- [laughs] -- Fischer? Yes, it is Fischer --

assemble the Fischer Kit, uh first select the -- the necessary pieces. // [P] These include -- hm -- those aren't easy shapes to describe. [P] Two -- red -- axles -- uh-how do I describe these shapes? -- I guess I could call them square. Two red -- square axles. That should be enough to differentiate them.* //

Following Voss et al. (1984), we occasionally inferred problems or solutions that writers had not articulated. For example, if the writer said, "I'm going back to insert something about what the tie looks like after Step 1," the writer was stating a solution. The inferred problem would be "forgot to specify what tie looks like after Step 1." Inter-coder reliability for segmenting protocols into problem units was 90%.

Some protocol comments did not fall into problem-solving segments, for example, comments that described ongoing processes ("Now, I'm writing...now I'm reading what I said"). These were noted but not analyzed.

The problem-solving segments were pooled and categorized into one of two general kinds of problems: (a) problems dealing with process, or how to go about writing; (b) problems dealing with content, or what is being written. Content problems fell into the following classes:

---

*(Underlining indicates words that writer was reading or writing.)

1. Content problems dealing with the accuracy of task specifications: determining and verifying task dimensions (the required order of steps, the correct position of an assembly piece, etc.);

2. Content problems dealing with selecting information for a user: deciding what task information the user needs and what the user can infer;

3. Content problems dealing with the mapping between language and task information: finding the word or wording that precisely reflects a given task referent or dimension (the word that corresponds most closely to a given shape, etc.);

4. Content problems dealing with purely textual features (apart from reference): (a) style and tone, (b) readability (word familiarity, sentence structure), (c) text conventions (mechanics, cohesion), and (d) format (paragraphing, headings, lists vs. prose).

We determined problem categories based on (1) what previous research has revealed about the kinds of problems writers consider and (2) our theoretical assumptions. Thus, the distinction between process and content problems has been widely observed in writing protocols (Hayes & Flower, 1980). The classes of content problems came out of our assumptions about the kinds of knowledge writers call on to write instructions -- knowledge about the task, the user, and the text -- and the

theoretical order in which they use that knowledge to formulate text: (1) they determine and verify task content from their task knowledge base, (2) they select aspects of task content for presentation in instructions, (3) they ensure those aspects are reflected precisely in the language of the instructions, (4) they ensure the language meets other criteria -- reads smoothly, complies with grammar rules, conveys the appropriate tone.

All writers articulated problems at each of these levels. But writers differed in the distribution of content problems across these four classes, in the types of problems posed in each class, in the methods of solution used, and in the evaluation criteria applied. We look at these differences separately for the tie-tying and assembly tasks.

Protocols of the Tie-tying Writers

Writers' problems. Table 10 shows what percentage of the total problems articulated by each writer fell into each class of problem. Relatively few problems dealt with process. Process problems were primarily of the sort: "I'll just make a note about that in the margin here and go back and add it in later" (Writer 11) or "I need to cross all this out because I recopied" (Writer 14). These segments revealed little that was not already apparent from writers' changes and other summary data. Significantly, problems dealing with higher-level, strategic planning were nearly absent. This finding reinforces the data on

TABLE 10. Percentage of Problems in Different Classes Posed by Each Writer (Tie-tying)

PROBLEM CLASS \ WRITER #	Most Effective 11	Most Effective 13	Least Effective 12	Least Effective 14
A. PROCESS	(8)	(15)	(7)	(6)
B. CONTENT				
1. Task Accuracy	(22)	(23)	(32)	(46)
a. Determine next task step a task specification	10	23	12	24
b. Verify task specification	9	0	18	11
c. Correct task error	1	0	0	11
d. Organize task steps	0	0	2	0
e. Assess applicability of task information	2	0	0	0
2. Information Selection	(37)	(31)	(22)	(8)
a. Assess information sufficiency for given task step	17	23	9	6
b. Assess whether step information is too detailed	9	0	5	0
c. Decide general level of detail for instructions	0	0	3	0
d. Choose hierarchical information (goals, etc.)	7	0	3	2
e. Choose feedback, recovery from error information	4	8	0	0
3. Language Mapping	(16)	(23)	(32)	(30)
a. Decide placement of global elements in text	2	0	2	2
b. Decide how to express specific task information	14	23	30	28
4. Textual Choices	(17)	(8)	(7)	(8)
TOTAL	100 (n=89)	100 (n=13)	100 (n=59)	100 (n=50)

surface measures, which showed that writers for the tie-tying task began to write almost immediately.

What concerns, if any, did writers express before writing? Ineffective Writer 14 said, "let me try this first," and ran through the tie-tying procedure briefly, globally describing his actions ("Now I pull it under...then over that end," etc.). Ineffective Writer 12 wondered, "should I organize first?" and decided not to because "the structure of the thing orders itself" (referring to the fixed order of tie-tying). Effective Writer 11 chose terms for the main task objects -- "The first thing I thought about...is how to describe the two ends of the tie and I've decided to call one end the fat end and the other end the skinny end."

These observations suggest:

1. Posing planning problems -- at least the type of planning problems seen in these writers -- did not seem to matter in writing instructions for this task. One effective writer (13) did not pose such problems, while two ineffective writers (12 and 14) did. This finding thus fails to confirm one of our basic predictions about effective writers.

2. How planning problems were solved and applied might have mattered. Thus, based on their comments, Writers 12 and 14 formulated superficially good planning strategies. But both writers produced overly segmented instructions.

Writer 14 went on to describe the task atomistically, even after analyzing it globally. His initial run-through seems to have served as a task accuracy check, not as a means of planning content. Writer 12 considered doing a preliminary task analysis but decided it wasn't necessary. We might hypothesize that writing a task outline that segments operations and defines end-states, then using this outline to help select information while composing, could aid writers who tend to over-segment and get lost in details.

Turning to content problems, we can look at differences between effective and ineffective writers in terms of the classes of problems shown in Table 10, focusing on the following: percentages of task accuracy vs. information selection problems, types of accuracy and selection problems, percentage of language mapping problems, and percentage of textual choice problems.

Accuracy vs. selection problems. An interesting difference seen in Table 10 is in the percentage of problems concerned with accuracy of task content (Class 1) as opposed to selection of task content (Class 2). Problem segments in the first category typically started out:

> "Now what's the next thing you do after crossing the tie over?" or "Where does my left hand go?"

Problem segments in the second category typically started out:

"Should I tell them which hand to use?" or "I need to say what the tie looks like at this point."

The distribution of problems in these two categories tended to distinguish both ineffective writers from both effective writers, and clearly differentiated ineffective Writer 14 from the rest. Table 10 shows that 46% of Writer 14's problem segments dealt with accuracy of task content, while only 8% of them dealt with selecting information to present to a user. For the other three writers, at least 22% of problems dealt with selecting information.

Why did Writer 14 invest more in determining and verifying the content of the instructions than he did in deciding how much content to present? What does this difference suggest?

Concern with how much information to present reflects a more rhetorical posture than concern with whether the information is correct. The writer is asking: Do users need more information? Less? Writer 14's lack of explicit concern with these questions may signal a general inattention to the needs of the instruction user. His overriding concern with accuracy suggests that he viewed writing instructions as a problem of task translation -- getting all the information precise and correct -- rather than as a problem of communication. Communication is necessarily selective, and Writer 14 did not worry about selection. This result supports our basic prediction that ineffective writers would pose fewer rhetorically based problems.

Writer 12 struck a better balance between accuracy and selection concerns. This suggests that Writer 12 was posing the same general category of rhetorical problems as the more effective Writers 11 and 13. Yet Writer 12's instructions were equally as over-segmented and ineffective as Writer 14's. Clues to what distinguished Writer 12 from the more effective writers may lie in a closer examination of the types of accuracy and selection problems writers talked about.

Types of accuracy and selection problems. Writers differed in the types of problems they considered in each of these two classes. As shown in Table 10, the most prevalent types of problems dealing with task accuracy included:

- Infomation-seeking problems -- determining the next task step ("what's the next thing you do?"); determining for a particular task feature, what the correct specification was ("which hand do you use?"). Writers solved these problems by executing the procedure and observing it.

- Information-verification problems -- verifying whether a piece of information, a step, or a constraint was correct. Writers solved the problem by executing the procedure.

Writer 13, a sprinter, posed no verification problems. This may reflect his decision to use a general level of description rather

than a detailed level. Thus, he dealt with fewer details such as instruments and measures that would warrant verification.

Two writers posed unique problems:

- Organizational problems -- Ineffective Writer 12 inquired whether the tie-tying procedure needed to "be structured." He decided that it didn't because its order was fixed.

- Application problems -- Effective Writer 11 inquired whether certain task information, e.g., giving hand specifications, would apply to the universe of users. He solved these problems by trying to anticipate other task environments (e.g., he represented left-handed users).

The task step organization problem posed by Writer 12 is one we expected to characterize effective writers. Concern with organization reflects an ability to achieve distance from the topic and to view the writing task from a top-down perspective. But in a linear procedure like tie-tying, explicit organizational concerns may not contribute to effective instructions. Writer 12, who was ineffective, was the only writer who expressed such concerns.

Writer 11 also posed problems that signal effective processes: problems that concern the applicability of task information to different users. Such problems assume a bird's-eye view, a global awareness of users.

These findings support one of our original predictions, that effective writers would consider more rhetorical or user-based problems (Writer 11). They do not support our prediction that effective writers exclusively would define problems at superordinate levels (the organizational question of Writer 12).

As shown in Table 10, there was considerable variation in problems dealing with selecting information. Drawing on our basic predictions about problem-solving in effective writers, we formed specific hypotheses about the problems in this class.

Our first hypothesis was that effective writers, as opposed to ineffective writers, would ask whether there was too much detail in a specification. This hypothesis comes from the prediction that effective writers would define more rhetorically based problems. We assume that writers who question the level of detail are aware that users can be confused -- they approach instruction writing as a problem of communication. This hypothesis was partially supported.

Writer 13, who was effective, did not consider problems of this type. A sprinter, Writer 13, wrote the leanest set of instructions of all four writers. The question of too much detail may not have been relevant to this method of writing.

The other three writers (11, 12, and 14) were plodders and produced longer and more detailed instructions than Writer 13. Questions about too much detail may be critical for such writers when they deal with continuous procedures like tie-tying.

Confirming these expectations, effective Writer 11 did articulate problems about excessive detail and ineffective Writer 14 did not. But ineffective Writer 12 did consider this problem, even as he produced overly detailed instructions. This observation reflects a rather consistent pattern and suggests there may be two levels of ineffectiveness. Writer 14 did not pose the problem of too much detail, thus neglecting a fundamental rhetorical question. Writer 12 posed the problem, but appears to have solved it ineffectively.

Our second hypothesis was that effective writers would consider more problems about including hierarchical and other global information (goals, structural and functional explanations, whole-task constraints). This hypothesis comes from the prediction that effective writers would define problems at more abstract levels. Prior findings suggest that better writers are more concerned with higher-level aspects of topic and text (Scardamalia, 1984; Flower, 1984). This hypothesis was not supported.

Effective Writer 13 articulated no problems of this type. Ineffective Writer 14 considered the addition and placement of whole-task constraints (for example, keeping the wide end of the tie face forward). Ineffective Writer 12 considered providing a general framework and orientation. Effective Writer 11 considered whole-task constraints as well as structural explanations. Thus, paralleling our findings from other sources

of data, writers who dealt with higher-level information did not necessarily produce effective instructions.

Our final hypothesis about information selection problems was that effective writers would consider more problems of providing feedback and recovery-from-error information. This hypothesis comes from our prediction linking effective writers with rhetorically based problems. We assume that writers who go beyond the essential task specifiers to consider feedback and error recovery are particularly sensitive to users' needs. By conceiving of users' error possibilities, they achieve distance from their own task knowledge and can transform that knowledge for a user (cf. Scardamalia, 1984). This hypothesis was supported.

As Table 10 shows, effective Writers 11 and 13 articulated problems of this type. In addition, both writers included feedback and recovery information in their texts. For example, Writer 11 included three major feedback checks ("Here's what it should look like now") in addition to simple end-states. Both writers included a major recovery loop at the end of the instructions ("If the narrow end is too long, undo the tie and start over"). Feedback and recovery information were not included in the ineffective writers' instructions, nor mentioned in their protocols.

Language mapping problems. Mapping problems arise when the writer has selected categories of information and wonders how to express them: "I wanna tell them that the wide end goes under,

what's a good word?" "How to describe this opening at the neck?" We expected that an effective writer would have more concerns of this sort, since they seem to reflect rhetorical distance from the text and awareness of alternative possibilities.

However, as Table 10 shows, ineffective Writers 12 and 14 articulated proportionately more mapping problems than did effective Writers 11 and 13. The critical distinction lay in the second type of problem in this class: problems concerned with how to express specific task information, such as an action, location, or orientation. These problems ranged from general -- "How do I convey this?" (referring to a given task dimension) -- to specific -- "Is it <u>bring</u> or <u>pull</u>?" (referring to action of moving the wide end of the tie). Besides action terms, specific problems dealt with terms for orientation ("should I say facing <u>out</u> or facing <u>up</u>?"), for instruments ("thumb <u>next to</u> index finger -- no, <u>against</u> -- I guess it actually is <u>against</u>"), and for constructed objects ("there is no word for this opening around the neck -- is it a hole?...a loop?"). To solve these problems, writers would typically execute the part of the task in question. They appeared to check the proposed terms against the actions to determine the best fit.

While users' performance could be helped or hurt by the terms used for constructed objects, our impression from watching users is that they were not affected by the particular words writers chose for actions and instruments. Thus, problems at this level reflect concern with referential precision that does

not appear to match users' needs. It was the two ineffective writers who were primarily concerned with subtle semantic distinctions. The thrust of these writers' comments ("I guess it actually is 'against'," "it seems to be...'bring' not 'pull',") suggests that they were seeking words that most accurately translated task details. Their comments seem to reflect a tendency toward preciosity and an approach to the writing problem as task translation.

Textual choice problems. The prevalent types of textual problems (considered by at least three writers) included (1) maintaining terminological consistency ("What did I call it before "[this end of the tie]?) and distinctions ("I already used 'outer' to refer to the loop"), (2) achieving simple language ("I should break this into two sentences," "use 'crossing point' instead of 'vertex', that's simpler"), and (3) using appropriate grammatical elements ("need to put this note in parentheses...").

Writer 11 voiced proportionately over twice as many textual problems as did other writers. In addition to the three categories above, he considered problems of tone and style ("need to use creative language here...but not too colorful"; "stick in a little word there to sound more instruction-like"), formatting ("I should be using...bullets and step numbers but I'm trying to just get it all written out"), and (7) use of motivators ("better say something to make them feel better -- maybe 'good luck')".

This textual orientation was not evident in effective Writer 13, a result that parallels the change data. Some instruction writers appeared to pay more attention to style and readability than others, but this did not influence their effectiveness.

The preceding discussion has focused on the kinds of problems writers talked about. We saw that, in contrast to the effective writers, the ineffective writers posed proportionately fewer kinds of problems that suggest rhetorical goals and more that suggest goals of task translation. But one ineffective writer (Writer 12) posed frequent rhetorical problems in some categories. Our next question, then, was whether he solved these problems differently from the effective writers. We look now at differences in how writers solved problems and evaluated solutions.

Writers' solutions and evaluation criteria. We predicted that effective writers would employ more rhetorical criteria -- those based on knowledge about users -- to evaluate solutions to the problems they posed. We will describe some of the differences we observed in solutions and criteria, focusing on accuracy and selection problems.

Solutions to accuracy problems were uniformly task-referenced, with one exception. In problems of applicability of information, as we saw earlier, effective Writer 11 referred solutions to his representation of a left-handed user -- a kind of internalized user testing. We see this method of evaluation as a potential prototype for effective writing.

Solutions to information selection problems are of more interest since we assume that user-based criteria are critical in deciding how much information to present. In fact, user-based criteria were applied frequently in selection problems. The difference between effective and ineffective writers lay not so much in whether they referred to a user, but in how flexibly they did so. Consider the problem of too much detail. Writers 11 and 12, effective and ineffective, both dealt with this problem, but they differed in solutions and criteria:

(1) Writer 12 always answered "no" to the question "Is this too much information?" or "Is this unnecessary?" Writer 11 varied -- he sometimes decided that the information was needed, sometimes that it was inferrable or too confusing.

(2) Writer 12 evaluated his solutions in terms of a fixed rule, one developed early in the writing process and from which he never varied: "There must be infinitely many ways they [users] can do this, like where they hold their hands, and not end up with a tie...So you have to keep narrowing their options." He referred to this rule whenever he posed questions about excessive detail: "It still gives them too many options unless I add...." By contrast, Writer 11 appeared to engage in internalized user testing to evaluate details at each choice point. This gave him more flexibility, and he arrived at different solutions depending on the results of the test. For example, having raised the question, "Is this too much?" Writer 11 reread the segment in question and decided, "[That should be

obvious]...I don't need to specify the obvious"; and at another point, "I don't want to blow this task out of proportion for them"; but at a third point, "I better be specific here since it's for someone who's never done this before."

Problems of information sufficiency showed a similar pattern. The criteria voiced by the two effective writers seemed to involve continually anticipating a user. Writer 13 made comments like:

> "They won't understand unless I tell them it's from the back" (referring to movement of tie). "It [the current level of specification] doesn't explain what I want them to do."

Writer 11 made comments like:

> "If they don't get this [specification], they might do it wrong and the tie wouldn't look good." "For clarity I'm gonna add...." "I need to trust their intelligence [in being able to infer this information]." "On second thought [after rereading a segment in question], no ambiguity is possible."

These comments suggest that both effective writers actively represented a user's responses in each context where they had questions about informational sufficiency. Writer 11 occasionally referred to other criteria, for example, the need for succinct language. But when these criteria conflicted with rhetorical goals, he put the rhetorical goals first. In one

case, Writer 11 shifted both criteria and solutions within a given problem segment:

> <u>The</u> -- I don't think I'm gonna need to write out the explanation of changing the tie to hand, I can kinda condense it to one sentence...<u>The ends of the tie should</u>...well actually, while I'm writing this I'm thinking that maybe I'll, since, I should be real specific since this is for somebody who's never done this...[deletes `ends...should'] <u>The fat end of the tie</u> [writes] <u>will be transferred</u>...that sounds good...[writes]...<u>to the right hand</u>...and I'll put a semicolon in...<u>and the skinny end will be transferred to the left-side</u> -- oops <u>the left hand</u>...

Here, the writer's initial decision not to elaborate information may be based on a kind of text esthetic ("condense it"), one that he expressed at other points in his protocol (e.g., "need condensed languge here"). But in reconsidering his decision, the writer subordinated the text-based rule to the anticipated needs of the user. This writer's ability to shift criteria within a problem-solving episode suggests an underlying flexibility of approach that may typify effective writers (Flower, 1984).

By contrast, ineffective Writer 12 invoked his option-narrowing rule when dealing with problems of information sufficiency: "there are still lots of ways they could do this"; "...haven't told them exactly where the fingers go." Thus, Writer 12 referred to the user, but these references appeared more rule-governed than rhetorical.

Although initially based on assumptions about the user, Writer 12's rule was ineffective for two reasons: First, it arose from a misconception about users -- not granting them the

ability to make inferences. Second, the rule was applied rigidly, in lieu of active user testing of information. We infer that Writer 12 was aware of his rule, and made a deliberate decision to use it, because of comments like the following. After the writer went back and orally summarized an operation he had just written out, he said: "That's general terms for what I just described in pieces."

We will briefly describe the other ineffective writer here (Writer 14), although he seldom posed information selection problems. When he did pose such problems, he did not mention the user, except for occasional implicit references ("That's confusing even to me"). In general, Writer 14 appeared to apply the same criteria for selecting information as for assessing task accuracy: "I should also mention that the left hand [stays steady]," "Somewhere I should say that the direction is to the right," "It just occurred to me that the right hand does all the movement," (for information just added:) "that's true." The principal criterion might be paraphrased as, "if you think of some information and it's an accurate characterization of the task, then include it." This reflects an essentially arhetorical perspective: the goal is to capture the task in all its correct physical nuance, regardless of the user's needs.

This writer's neglect of rhetorical criteria is especially clear in the following problem-solving segment, which occurred during his after-draft review. He had just read his over-segmented description of one of the critical task

operations: "That's confusing even to me--[rereads the sentence twice, trying to execute the instruction, finally coming up with the correct actions]...oh, I see...yes, that's right." The writer initially recognized users' potential confusion, but then ignored this potential after he figured out what he was trying to say. Regardless of the complexity of his description, it exactly captured the task detail this writer aimed for.

Therefore, Writers 12 and 14 appeared to arrive at ineffective, overly segmented instructions for different reasons: Writer 14 because he failed to apply user-based criteria, Writer 12 because he applied them in the form of fixed rules, not in the form of flexible, context-specific user testing.

Summary. Based on the protocol comments, we can infer that writers' effectiveness in the tie-tying task lies primarily in how they dealt with selecting information to be presented in instructions. As a group, the effective writers considered information selection problems more than the ineffective writers. Since selection could be viewed as a user-centered question, this supports our prediction that effective writers would pose rhetorically based problems. Moreover, effective writers appeared to evaluate solutions by anticipating a user's response to specific choices of information. This supports the prediction that effective writers would apply rhetorically based criteria.

We found no support for the prediction that effective writers would define more abstract problems or use a top-down approach to composing instructions. Nor was there evidence that

effective writers posed planning problems. Indeed, all writers appeared to follow a narrative organization in retrieving and applying their procedural task knowledge.

We found two levels of ineffective writers. One ineffective writer rarely posed selection problems. The other ineffective writer posed selection problems, and applied what appeared on the surface to be rhetorical criteria. But these criteria did not seem to involve active user representation. Instead, the writer developed a general user-based rule to reduce the procedure to fine detail, a rule he then applied inflexibly.. Both ineffective writers appeared to approach the writing problem as precise task translation -- they operated as "knowledge-tellers," in the terms of Scardamalia (1984).

These distinctions in how writers defined instruction writing emerged vividly in their approach to review. One effective writer (13) stated that he was reviewing "to see if it read lucidly," and predicted "it [would] get muddy" in the description of the knot-forming operations. The other effective writer (11) reread to "see if it makes sense." Effective writers may have seen review as a chance to adapt the text to the user's needs for information.

By contrast, one ineffective writer (12) did not review because, as he put it later, he thought he had captured as much detail as he could. The other ineffective writer (14) reviewed by rewriting part of his draft. When this writer encountered a complicated description, he detected the complexity but decided

to keep it, trading clarity for precision. Moreover, the third ineffective writer (4), who did not think aloud, said during debriefing that he reviewed to "be sure that right and left hand were OK...that I hadn't changed hands and forgot to mention it." Ineffective writers may have defined review as a chance to assess task accuracy.

The nature of the tie-tying procedure -- short, strictly ordered, and over-learned by the writer -- may have militated against planning and defining abstract problem hierarchies. We can compare these findings with those for the assembly procedure -- longer, more complex, and with more options for ordering.

Protocols of the Assembly Writers

We collected protocol data from Writers 2, 3, 6, and 9. Writers 3 and 9 produced the least effective assembly instructions and contrast clearly with the other six assembly writers. Writers 2 and 6 produced instructions of medium effectiveness. Therefore, while we cannot take their protocol comments as evidence of exemplary problem-solving processes, we can look at aspects that contrast with those of the two ineffective writers.

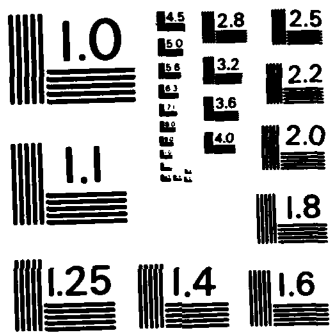
Writers' problems. Table 11 shows what percentage of the problems posed by each writer fell into each major class of problems. The distribution of problems across the major classes does not clearly distinguish the more effective writers (2, 6) from the less effective writers (3, 9). Our expectation based on the tie-tying group was that ineffective writers would have lower proportions of problems dealing with selecting information, but this was not clearly borne out. Therefore, after looking briefly at process problems, we will focus on specific types of problems writers posed in each class and how they solved them.

Writers posed relatively few process problems. But, unlike the tie-tying group, writers did pose process problems that dealt with higher-level planning. These problems occurred in one effective writer (6) and one ineffective writer (3), and they concerned developing a global outline. Counter to our original

AD-A152 083 PROCESSES INVOLVED IN WRITING EFFECTIVE PROCEDURAL INSTRUCTIONS(U) AMERICAN INSTITUTES FOR RESEARCH WASHINGTON DC V M HOLLAND ET AL. 28 FEB 85 2/2

UNCLASSIFIED N00014-83-C-0590 F/G 5/7 NL

END



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

TABLE 11. Percentage of Problems in Different Classes Posed by Each Writer (Assembly)

PROBLEM CLASS \ WRITER #	Moderately Effective		Least Effective	
	6	2	3	9
A. PROCESS	(6)	(6)	(7)	(5)
B. CONTENT				
1. Task Accuracy	(7)	(22)	(25)	(20)
a. Determine next task step	0	12	8	11
b. Determine or verify specific information category	6	10	13	7
c. Correct task error	1	0	4	2
2. Information Selection	(48)	(60)	(36)	(41)
a. Choose what piece goes first	1	2	2	2
b. Assess information sufficiency for selection or correction operation	27	40	22	32
c. Choose goal and hierarchical information	1	2	6	7
d. Choose feedback	6	10	0	0
e. Order task steps				
o overall	9	0	0	0
o local	1	6	0	0
f. Group task steps into chunks	4	0	6	0
3. Language Mapping	(11)	(6)	(19)	(18)
a. Decide how name piece	6	6	4	7
b. Decide how to express specific information	5	0	15	11
4. Textual Choices	(28)	(6)	(13)	(16)
a. Decide sentence-level features	16	6	7	16
b. Decide global features (headings, etc.)	12	0	6	0
TOTAL	100 (n=108)	100 (n=50)	100 (n=53)	100 (n=44)

predictions, we could not distinguish effective from ineffective writers on this dimension.

Types of accuracy and selection problems. Effective and ineffective writers were not readily distinguishable in terms of the types of task accuracy problems they posed. Writer 6 stands out as having posed no problems dealing with determining the next task step, while the other three writers posed relatively substantial proportions of such problems. The reason for this difference is that Writer 6 used a preliminary outline to determine the sequence of task steps. She derived the sequence in her outline by comparing several alternative orderings and selecting one. Thus, we classified her task step ordering problems as information selection problems.

Given the moderate proportions of the other writers' problems dealing with "determining the next task step," we can infer that these writers frequently used a "what's next" strategy to determine task ordering. They appeared to treat ordering not as a problem of choice, but as one of retrieving operations from a linear task schema ("What's the next thing to connect?...Oh yes, the steering wheel."). Since the "what's next" group included one effective as well as two ineffective writers, these results provide inconsistent support for the hypothesis that better writers would define more abstract, hierarchically structured problems.

Looking at types of selection problems, we find three differences of interest. First, corresponding to her lack of "what's next" problems in the accuracy class, we find that more effective Writer 6 had a proportionately higher percentage of problems dealing with ordering the task steps: "There are lots of ways to sequence these steps, which should I go with?" In addition, Writer 2, also more effective, posed problems of this type. Thus, Writer 2 decided task sequence sometimes in terms of "what's next," sometimes in terms of a superordinate representation of alternatives. The two ineffective writers posed no alternative ordering problems.

However, the two more effective writers differed in the degree to which they considered alternative ordering, as Table 11 shows. Writer 2 formulated ordering possibilities only at selected local choice points, such as within a subassembly, ("Which of the three steering pieces should go first? I guess I'll give them the easy one..."). He did not deal with overall task ordering. Considering other features of this writer's composing process (he did not do higher-level planning; he embarked almost immediately upon drafting; the majority of his questions were of the "what's next?" variety), his general approach to ordering appears basically linear.

Writer 6 used a hierarchical approach. She planned extensively and dealt with overall task ordering. Thus, while concern for alternative orders may be associated with writing effective instructions, the more hierarchical approach to the

ordering problem was not noticeably better than the more linear approach.

A second difference of interest was in problems of providing feedback. The two more effective writers talked about providing feedback and included feedback in their texts (for example, "You should now be able to pull the vehicle along by the trailer piece.") The two ineffective writers did not talk about feedback and included little of it in their texts. This result replicates the findings from the tie-tying writers and supports the prediction that effective writers would consider more user-based problems.

A final difference of interest was in problems of providing goal statements. Ineffective writers articulated more of these problems. This result is counterintuitive, especially if we refer back to the number of goal statements in each assembly writer's instructions (Table 5). Apparently, the more effective Writers 2 and 6 wrote numerous goal statements that they did not talk about. Because they composed these statements while they wrote, we infer that including goals was at some level automatic for them. Ineffective Writers 3 and 9 talked about these statements more because they added them as deliberate, delayed changes during or after drafting. The integration strategy of Writers 2 and 6 may be related to their effectiveness: it is conceivably better to write goals as part of the initial specification of an operation, because then the writer can take immediate advantage of the operational constraints goals impose.

Mapping problems. As in the tie-tying writers, we find that the ineffective writers posed somewhat higher proportions of problems dealing with how language maps to task. Many of these problems concerned subtle choices: "Do you push the peg in or press it in? I guess I'm really pressing...." In most cases, subtle semantic differences did not appear to affect users' performance. These writers' concerns with mapping may have arisen from arhetorical goals of precise task translation or from rhetorical goals in which users' primary needs were misconceived.

Textual choice problems. Problems in this class do not differentiate more and less effective writers. We briefly mention the findings because they replicate those from the tie-tying writers. More effective Writer 6 posed about twice as many textual problems as did either ineffective writer. She was concerned with both local textual matters (familiar words, simple sentences) and global design matters (parallel sentences and paragraphs). More effective Writer 2, on the other hand, posed half of many textual problems as did either ineffective writer. Thus, some writers appeared more style-oriented than others, but this did not bear on how effectively they wrote instructions.

Writers' choice points, solution methods, and evaluation criteria. In looking at the types of problems posed by more vs. less effective writers, we were particularly interested in any differences that might surround the most prevalent type of information selection problem: assessing the sufficiency of information about task operations. Problems of this type

frequently began, "Have I told them enough about this piece (or how to do this)?" We speculated that this type of problem would be importantly tied to the informational adequacy of the instructions. However, all four writers posed problems of this general type. Therefore, we looked for possible differences in how writers dealt with informational sufficiency problems: where writers posed these problems, how they solved them, and how they evaluated solutions.

While the points at which writers posed problems were not part of our analysis of problem segments, we discuss this dimension because it emerged clearly when we read the protocols. The two ineffective writers did not consistently pose problems about informational sufficiency at what we deemed critical points of choice. For example, in deciding what to name a piece, all four writers occasionally used technically accurate names like "grommet," "wingnut," "axle," and "hub." They did not question whether users would know these names. However, users could find the piece as long as the writer provided sufficient physical descriptors. It is significant, then, the ineffective writers failed to ask, "What else do I need to describe this piece?" at many points where they had used technical names. The two more effective writers always considered "what else?" when they used a technical name. As a result, their instructions were comparatively long and redundant while those of the ineffective writers were comparatively lean. Users who read the long, redundant instructions usually picked out the right pieces, while users of the lean instructions often picked out wrong pieces.

More and less effective writers also used different methods of solving information sufficiency problems. We can compare ineffective Writer 9 with more effective Writer 2 in terms of how they dealt with the problem of describing pieces. Writer 2 used a general solution before he began to write. He grouped the pieces into similarity sets and inventoried features that distinguished the sets and distinguished pieces within sets ("Here's two of those...four round ones...").

Writer 9's method of solution was less systematic. He did not inventory distinctive piece features before writing a description. He would, after writing, frequently scan the array of pieces to see whether a descriptor uniquely distinguished the target piece ("Is that the only one that's round?").

Finally, more and less effective writers differed in their evaluation criteria for problems of informational sufficiency. When evaluating a description he had written, Writer 2 often added extra descriptors. At one point he reasoned, "...need to describe it clearly, since they have to pick it out without pictures." He occasionally articulated a conflicting criterion, that of the need to be concise ("It's getting long," "awful lot of words to describe two little pieces"). But the redundancy criterion took precedence. This ordering of criteria appeared to stem from the writer's understanding of the user's needs for information given the constraint of no pictures.

Compare how Writer 9 evaluated a description he had written:

> pick out four wheels. I'm assuming they--they know the difference between wheels and tires, although...let's see here. Um. Four wheels--the one with the wide [deletes "wide"]--The four red wheels, since they--since they're the only wheels, that should do--No--they're not the only wheels--there are these too. With a--wide, flat surface. OK, so that's something that'll differentiate those wheels from the others in the kit.

In this segment, the writer considered deleting the extra descriptor "wide" because he thought "four wheels" uniquely specified the pieces. He added "wide" only when he discovered another set of four wheels. For problems of this sort, Writer 9 appeared to follow this rule: "If there are enough descriptors to distinguish a piece uniquely, don't add any more."

What goals motivated this rule? Writer 9's underlying goals in piece description and other selection problems seemed to be attaining an elegant, smart-sounding set of instructions. He would comment, "Don't spell it all out..." or "sounds too simple-minded," when he decided not to elaborate information. The following segment is also illustrative:

> What's the term for that..."wrench"?...no, "center"? -- I'll have to call it just "center." Lousy terminology. (The writer later remembered the term he wanted and changed "center" to "rotating wingnut.")

Again, the concern seemed to be with achieving technically accurate terms, a text-centered goal applied at the expense of clarity.

In evaluating his minimal piece descriptions, Writer 9 sometimes referred to user-based criteria -- "they should be able to pick it out from that," etc. But these criteria did not appear to involve actively testing a user's response. They were more like generic rationales, subordinate to the text esthetic. Similarly, after completing his draft, Writer 9 decided against making more than minor informational additions or changes because "it might make it more confusing by giving them more instructions than they need." In this criterion, the writer appreciated the possibility of over-sufficiency, but he did not weigh the possibility of insufficient specification.

These results suggest that, in dealing with problems of information sufficiency, effective and ineffective instruction writers may differ in their ability to identify choice points that are critical in terms of users' needs for information. It is at these points that effective writers appeared to focus their user-based questions. Writers may also differ in whether they use systematic methods to scan the kinds of information available to users. Finally, writers may differ in the flexibility of their user-based criteria: e.g., in whether they anticipate the user's response with respect to both the possibility of too much information and the possibility of too little.

Looking at other kinds of information selection problems, we found further signs that ineffective writers applied rhetorical criteria, but differed in the points at which they applied them and in the flexibility with which they conceived them. For

example, in determining how to divide the instructions into chunks (after drafting), ineffective Writer 3 explicitly played the role of user:

> OK, now---how can I break these down into little groups? OK. So the task isn't overwhelming for them. Now, I'm sitting in front of this group--you oughta understand what I'm doing. I'm looking at these--group of objects, and I should know what I'm trying to do with them.

This writer articulated numerous user-based criteria but applied these at levels that apparently were not critical to effectiveness: deciding how and where to number steps ("so it's easier to follow"), deciding how and where to place headings ("so the task isn't overwhelming").

Ineffective writers also gave none of the signs of applying flexible user representations that we sometimes saw in the more effective writers. Flexible representations permit testing the rhetorical consequences of writing choices. For example, Writer 6 evaluated two organizational strategies -- to have the user select all the pieces first, or to have the user select each piece immediately before the relevant connection operation:

> ...maybe I should say the first step is to pick out all the pieces you want -- although that seems kind of artificial. Oh dear. 'Cause I don't think most people -- I think most people would just pick the pieces right out of the kit when they needed them. And that would be the thing I would naturally do, but -- hm. But if I have them pick up the pieces as they go along -- there's gonna be all these little extra instructions along the way. I don't know. Hm. Well -- let's see, if I have them pick up the pieces as they go along, they're going to have all these steps saying "Now find the piece that's this and now find the piece that's that." But that might not be so bad. The other thing is,

> if I define the piece right there, it might help them to identify the pieces -- that they've selected the right one -- if they are trying to do something with it. 'Cause if they picked the wrong one and then found it didn't fit in the slot or something like that, that would be one way to know you had the wrong piece. Well, I think I'll try having them find the pieces as they go along. I'll try it for the first couple of steps and see how it goes.

This writer built a rich representation of users that allowed her to anticipate their ongoing response -- potential confusion, error, recovery -- to different ways of organizing the task. She decided on a particular organization based on how she expected users to respond.

Summary. The protocol data on assembly writers provided limited support for some of our original hypotheses and led to new hypotheses.

Did better writers of assembly instructions define more abstract, hierarchically structured problems? To some extent they did. They tended to represent task-ordering alternatives. But equally effective writers used either a linear approach, in which they considered selected local alternatives while writing, or a hierarchical approach, in which they considered alternatives for the whole task before writing.

Did better writers consistently define higher-level planning problems? They did not. Global planning characterized one effective and one ineffective writer. But the ineffective writer planned between the first and second drafts and dealt not with selecting content, but with chunking the text as given under

superordinate headings. Moreover, one effective writer engaged in a kind of local planning: inventorying features of assembly pieces before writing decriptions.

We can hypothesize that the key to writing effective instructions is not the degree to which strategies like planning and posing alternatives are part of an abstract, hierarachically structured problem definition. Rather, the key is using those strategies to guide information selection, applying them at points and levels that are critical to a user's needs for information.

Did better writers pose more rhetorically based problems and apply rhetorically based criteria? In general, the results suggest not. Ineffective writers, for example, considered problems of selecting information, providing goals, and formatting the text into manageable chunks. They frequently referred to users' needs.

However, ineffective writers appeared to differ from more effective writers in the points at which they applied user-based problems and criteria and in the flexibility of these criteria. Thus, ineffective writers did not identify as choice points certain critical points in selecting information -- e.g., after the introduction of inscrutable piece labels. They frequently focused rhetorical strategies on higher levels (e.g., global text arrangement) or lower levels (e.g., precisely mapping word to referent) rather than on basic-level, task step information. A

question for further study, then, is how writers identify choice points as they compose of a set of instructions.

Moreover, ineffective writers tended to apply user-based criteria generically rather than anticipating users' response to specific informational choices. When they posited text-based criteria, such as "be concise," they did not typically weigh these against user-based criteria.

These findings reflect those for writers in the tie-tying group. User-based criteria may be ineffective if they are applied inflexibly and without adjustment to context.

In both groups, tie-tying and assembly, we could interpret the differences in how more and less effective writers treated evaluation criteria in terms of how they represented communication meta-goals. These goals arise from basic assumptions about communication, such as defined in Grice's conversational maxims (Grice, 1975). Two competing goals that must be balanced are defined by Grice's maxim of quantity: "be informative" and "don't be more informative than required." Ineffective assembly writers emphasized the latter. Ineffective tie-tying writers emphasized the former. Effective writers tended to weigh both sides of the maxim, and to consider the consequences to the user of both redundant and concise specifications.

IV. CONCLUSIONS

Summary

The key distinction between effective and ineffective instructions for two kinds of procedural tasks lay in the sufficiency of the information used to specify the task steps. Thus, ineffective instructions tended to leave critical information categories unspecified -- end-states or directional specifiers for the movements of tie-tying; for assembly, descriptions, locations, or orientations of connecting pieces. Ineffective instructions did not appear to be distinguished by consistent stylistic features -- such as longer, more convoluted sentences or obscure vocabulary. Hierarchical features (goal statements, explanations) distinguished effective instructions primarily to the extent that those features added informational constraints.

Similarly, the key distinction between writers of effective and ineffective instructions lay in how they dealt with selecting information from their base of task knowledge. Thus, we did not find consistent differences on many of the expected composing dimensions: planning before writing, concern with global aspects of text, extensiveness of changes, changes to text meaning. In fact, the striking variety of composing styles we found observed spanned the effectiveness dimension. For example, textually

oriented writers -- who revised extensively, made numerous stylistic changes, and expressed proportionately more concerns with sentence structure, tone, and formatting -- were both effective and ineffective.

How did effective writers deal with selecting information while they composed? In general, we can infer that they:

- considered more problems about selecting information ("Did I tell them enough to understand this step?" "Did I tell them too much?") and proportionately fewer problems about accuracy of information or precise word meaning;

- applied selection problems at critical choice points in the emerging text, that is, critical in terms of whether users could understand the instructions;

- tested solutions to these problems by anticipating the effects on a user of informational choices;

  subordinated text-based criteria to user-based criteria when these conflicted;

- were concerned with providing feedback and recovery-from-error information in both tasks, and in the assembly task, with generating

alternative solutions (for example, for how to order task steps).

Effective writers also tended to use process strategies that helped them select and test information:

- For the assembly task, the most effective writer did a detailed pre-writing plan of the categories of content to be presented in each task step; other effective writers did more limited content planning, such as scanning and classifying the array of information available to users. Those who didn't write out a plan reviewed text while writing, in informational chunks of approximately step-size.

- For the tie-tying task, effective writers reviewed the whole text after writing, seeking to test information for user understanding. A whole-text review appeared to help writers evaluate information in terms of how they had segmented the tie-tying procedure.

How did ineffective writers deal with information selection? Some writers rarely considered problems of selection, focusing instead on accuracy and precision, and they rarely applied user-based criteria, calling instead on text- or task-based criteria. Other writers posed user-based selection problems but did so at points that were not critical to users' needs for

information. Moreover, they applied user-based criteria but did so inflexibly or by misunderstanding users' needs ("they need to have all their options narrowed"). These writers might define selection criteria on rhetorical grounds at the outset of writing, but turn these criteria into fixed rules. These rules would replace ongoing construction of a user and evaluation of information with respect to that construction.

Ineffective writers also failed to employ process strategies that would help them select and test information. For example, in the assembly task they did not plan task content before writing or review it in step-size chunks while drafting. In the tie-tying task, they failed to review the whole text, or failed to review it for clarity, or neglected to make changes when they detected ambiguity.

Why did ineffective writers fail to anticipate users' informational needs while writing? There are several possibilities:

1) They were not aware that active user representation was crucial. 2) They were aware but made a strategic decision to focus on the task rather than on the user. 3) They tried to represent the user but did it ineptly or inaccurately. Why did these writers tend to rely on rules rather than on flexible criteria? Appealing to rules may reduce the demands of continually constructing and reconstructing user responses.

Practical Implications

If focusing on the procedural task kept writers from attending to the user, then writing out a task plan first could help by reducing the writer's burden. Other possible sources of ineffective writing -- the writer was not actively aware of the user or misunderstood the user's needs -- would require interventions in writers' knowledge and rhetorical strategy. Our study of writers revising their instructions (Appendix B) showed that watching a user on videotape helped writers detect informational ambiguities in their texts. Showing writers a user who makes errors could be one way of sensitizing them to users' needs.

If writers develop a task plan to help them write better instructions, what should the plan consist of? Our results suggest that it should include a detailed task analysis that (1) lays out the order and boundaries of each task step, giving end-states, and (2) defines the categories of information critical to specifying a task step generally (e.g., location on three dimensions.)

Our results also suggest that this pre-writing plan should not be merely an instantiation of a text schema -- e.g., one that lays out categories like "introduction," "goal of task," "steps," "feedback." Effective writers, when they used a written plan, either started with a task-analytic outline or turned to one after trying a text schematic outline.

Producing a task plan could potentially solve several of the problems we observed in instruction writing: (1) easing the memory burden for step order and content, (2) ensuring that critical information categories are not left unspecified, (3) controlling level of detail, (4) reducing over-segmentation in continuous procedures like tying a tie.

A further process-based implication of our results is that instruction writing might be improved if writers initially shelve concerns with formatting, readability, precision of word choice, and text logic and esthetics (e.g., the ordering and subordination of headings). These concerns can be handled later by the writer, by an editor, or by an automated writer aiding system based on text features of particular interest (e.g., Kieras, 1984). Our results suggest that when time is limited, the writer's attention is best given to issues of information selection and adequacy.

Implementation

These recommendations about how to proceed in writing technical instructions are preliminary. We offer them as hypotheses. Eventually, tested recommendations could be operationalized and included in a process-based writer aiding system--one that applies at the beginning of the writing process. This system could be used together with product-based editing aids, which analyze completed text, such as the Writer's Workbench (Macdonald et al., 1982) or the system proposed by Kieras (1984).

Theoretical Questions

In examining the kinds of problems writers considered, we found clues to the nature of their writing goals and how these goals interacted and developed in the course of composing. We did not attempt a formal analysis of goals, but we noticed that:

- Goals frequently conflicted within domains (e.g., "write concisely" vs. "specify fully") and across domains (e.g., content vs. process goals--"that's not the right word, but I need to move on"). Effective writers appeared more likely to represent and articulate these conflicts.

- Goals assumed a certain hierarchy, in terms of what concerns could interrupt others and in terms of the kinds of meta-goals by which goal conflicts were resolved.

- Goals seemed to evolve in the course of writing, from more text-based or schema-based to more rhetorical (solutions became more flexible as a result) -- but evidence for this occurred only in effective writers' protocols.

Further work could explore goal interaction and development in technical writing, using more focused methodology. Concurrent thinking-aloud protocols vary within and between writers in the degree and depth to which writers articulate their criteria for

writing decisions. Protocols based on retrospective probes (e.g., Odell & Goswami, 1982) could be used to reveal how instruction writers represent and order specific goals and manage their trade-offs. Moreover, looking at writers' goals in terms of basic conversational principles (Grice, 1975) seems a potentially useful way to compare writers' approaches.

Limitations of Conclusions

Our observations about what makes instructions and writers effective are based on users' immediate performance of instructions rather than on their memory for instructions. Immediate performance was tested in a context where users were highly motivated. Mean differences in time to read were on the order of minutes. Contexts that emphasize memory, attitude, or reading speed in seconds might show more advantage for hierarchically structured instructions, clean formatting, simple sentences, etc. Moreover, the procedural tasks we looked at were relatively simple. Writers of instructions for longer, more complex procedures might do more higher level planning and reordering than the writers we observed.

References

Atlas, M. (1981). Addressing an audience: A study of expert-novice differences in writing (Technical Report). Carnegie-Mellon University: The Document Design Project.

Baggett, P. (November, 1983). Learning a procedure from multimedia instructions: The effects of film and practice. Technical Report #125, University of Colorado, Institute of Cognitive Science.

Baggett, P., & Ehrenfeucht, A. (1981). How an unfamiliar thing should be called. Technical Report No. 111-ONR, University of Colorado, Institute of Cognitive Science.

Beach, R. (1976). Self-evaluation of extensive revisors and nonrevisors. College Composition and Communication, 27, 160-164.

Beiger, G.R., & Glock, M.D. (1982). The information content of picture-text assembly instructions. Technical Report No. 5, Cornell University, Department of Education.

Bridwell, L.S. (1980). Revising strategies in twelfth grade students' transactional writing. Research in the Teaching of English, 14(3), 197-222.

Collins, A.L., & Gentner, D. (1980). A framework for a cognitive theory of writing. In L.W. Gregg & E. Steinberg (Eds.), Cognitive processes in writing. Hillsdale, NJ: Lawrence Erlbaum Associates.

Dixon, P. (1982). Plans and written directions for complex tasks. Journal of Verbal Learning and Verbal Behavior, 21, 70-84.

Duffy, T.M. Preparing technical manuals: Specifications and guidelines. Presented at American Educational Research Association, New York, 1982.

Faigley, L., & Witte, S. (1981). Analyzing revision. College Composition and Communication, 32, 400-414.

Felker, D.B., Pickering, F., Charrow, V., Holland, V.M., & Redish, G. (1981). Guidelines for document designers. American Institutes for Research, Document Design Project.

Fillmore, C.J. (1968). The case for case. In E. Bach and R. Harms (Eds.), Universals in Linguistic Theory. New York: Holt, Rinchart & Winston.

Flower, L. The expert's option: Conscious problem-solving during planning and revising. Paper presented at the AERA meeting, New Orleans, 1984.

Flower, L.S., & Hayes, J.R. (1980). Cognition of discovery: Defining a rhetorical problem. College Composition and Communication, 31(1), 23-32.

Flower, L.S., & Hayes, J.R. (1981). Plans and the cognitive process of writing. In C. Frederiksen, M. Whiteman, & J. Dominic (Eds.), Writing: The nature, development, and teaching of written communication. Hillsdale, NJ: Lawrence Erlbaum Associates.

General Accounting Office (1979). Improved management of maintenance manuals need in DoD. LCD-79-105.

Gordon, L., Munro, A., Rigney, J.W. & Lutz, K.A. (1978). Summaries and recalls for three types of text. Technical Report No. 85, University of Southern California.

Grice, H.P. (1975). Logic and conversation. In P. Cole & J.L. Morgan (Eds.), Syntax and semantics (Vol. 3). New York: Academic Press.

Halpin, J. Approaches to revision in experts and novices. Paper presented at the AERA meeting, New Orleans, 1984.

Hayes, J.R. Detecting text problems. Paper presented at the AERA meeting, New Orleans, 1984.

Hayes, J.R., & Flower, L.S. (1980). Identifying the organization of writing processes. In L.W. Gregg & E.R. Steinberg (Eds.), Cognitive processes in writing. Hillsdale, NJ: Lawrence Erlbaum Associates.

Holland, V.M. (1981). Psycholinguistic alternatives to readability formulas. Technical Report No. 12, American Institutes for Research, Document Design Project.

Houp, W., & Pearsall, T.E. (1980). Reporting technical information. Beverly Hills: Glencoe.

Jonassen, D.H. (1982). Individual differences and learning from text. In Jonassen, D.H. (Ed.), The technology of text. Englewood Cliffs, N.J.: Educational Technology Publications.

Kern, R.P., Sticht, T.G., Welty, D. & Hauke, R.N. (1976). Guidebook for the development of army training literature. Washington, D.C.: U.S. Army Research Institute for the Behavioral and Social Sciences, HumRRO.

Kieras, D.E. (1979). The role of global topics and sentence topics in the construction of passage macrostructure. Technical Report No. 6. Tucson, Ariz.: University of Arizona.

Kieras, D.E. (1984). Reading in order to operate equipment. Paper presented at the American Educational Research Association, New Orleans.

Kintsch, W. & van Dijk, T.A. (1978). Toward a model of text comprehension and production. Psychological Review, 85, 363-394.

Lutz, J. A study of revising and editing at the terminal. Unpublished paper, English Department, Miami University, Oxford, Ohio, 1983.

Macdonald, N.H., Frase, L.T., Gingrich, P.S., & Keenan, S.A. (1982). The writer's workbench: Computer aids for text analysis. IEEE Transactions, 30(1), 105-110.

Mandler, J.M., & Parker, R.E. (1976). Memory for descriptive and spatial information in complex pictures. Journal of Experimental Psychology: Human Learning and Memory, 2, 38-48.

Meyer, B.J.F. (1977). The structure of prose: Effects on learning and memory and implications for educational practice. In R.C. Anderson, R.J. Spiro, & W.E. Montague (Eds.), Schooling and the acquisition of knowledge. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Odell, L. & Goswami, D. (1982). Writing in a non-academic setting. Research in the Teaching of English, 16(3), 201-224.

Peterson, J. The composing processes of adult writers and revisers. Unpublished paper, American Institutes for Research, Washington, DC, 1983.

Reder, L.M., Charney, D.H., & Morgan, K.I. (1984). The role of elaboration in learning a skill from instructional text. Technical Report No. 1, Carnegie-Mellon University, Department of Psychology.

Scardamalia, M. Development of cognitive strategies in writing. Invited address, the AERA meeting, New Orleans, 1984.

Schorr, F.L., & Glock, M.D. (1983). Comprehending procedural instructions: The influence of comprehension monitoring strategies and instructional materials. Technical Report #10, Cornell University, College of Agriculture and Life Sciences.

Smith, E.E., & Goodman, L. (1982). Understanding instructions: The role of explanatory material. (TR-5088) Bolt Beranek and Newman, Inc.

Sommers, N. (1980). Revision strategies of student writers and experienced adult writers. College Composition and Communication, 31(4), 378-388.

Stone, D.E., & Glock, M.D. (1981). How do young adults read directions with and without pictures? Journal of Educational Psychology, 73, 419-426.

Voss, J.F., Green, T.R., Post, T.A., & Penner, B.C. (1983). Problem solving skill in the social sciences. In G.H. Bower (Ed.), The psychology of learning and motivation: Advances in research theory, 17, 165-213. New York: Academic Press.

APPENDIX A

SELECTED WRITERS' INSTRUCTIONS

ASSEMBLY: WRITER 15 - most effective

WRITER 16 - moderately effective

WRITER 3 - least effective

TIE-TYING: WRITER 13 - most effective

WRITER 4 - least effective (under-specified)

WRITER 12 - least effective (over-segmented)

Writer 15

A. Introducing the task

In this task, you will assemble a small toy car. It will have a steering wheel that turns the front wheels. It will also have a towing piece at the back.

B. Choosing the pieces you'll need

These are 36 pieces in the Fischer Kit, but for this task you'll need only 14. Spread out the pieces in front of you and select:

(1) the chassis -	a red rectangle, 3-1/2 by 1-3/4 inch with six notches down each long side and nine holes down the middle.
(2) the front axle -	a red rectangle, 3 by 1/2 inch, with three notches down each long side, a round hole in the center and two notched tubes sticking out at each end.
(3) the steering column -	a grey thin column, 2-1/8 inch long with a square peg at one end round at the other end.
(4) the steering column clamp -	a red C-shaped object, a little bigger than the eraser on a pencil.

Now gather all the round red plastic pieces. Two are larger than the others. You'll need one of each:

<u>The steering wheel</u>

(5) the steering wheel top -	flat on one side, a small hole through, with a screw end.
(6) the steering wheel bottom -	same size as steering wheel top, large hole in middle where top piece screws in.

(7) the small black rubber ring -	the smallest of the three black rubber rings, the one without "treads" or notches.
(8 & 9) the back axles	these look like half-inch red squares on a red stick - small square ping-pong paddles - with a square peg on one side and a long notch on the other side.
(10) the towing piece -	a red piece about 1/2 inch long - with a rounded end, two irregularly shaped holes (don't pick the piece with three holes).
(11, 12, 13, 14) - the wheels	pick the one with flat "treads" - you'll need <u>four</u>.

C. Putting the pieces together

(1) <u>Assemble the front-end</u>

a) Stick the round end of the steering column into the square hole in the center of the front axle.

b) Press the square peg securely into the square hole.

(2) <u>Attach the front-end to the chassis</u>

a) Stick the round end of the steering column through the first <u>full</u> hole on either side of the chassis. (Don't stick it through the three-quarter holes on the edge of the chassis.)

b) Push the front axle up until it is flush with face of the chassis.

c) Slide the steering column clamp down the steering column until the chassis and the front axle are tightly connected.

(3) <u>Assemble the steering wheel</u>

a) Hold the steering wheel top - the one with the screw in it - with one hand.

b) Put the black ring around it like a tire on a wheel.

c) Then take the steering wheel bottom - the piece with the wide hole in it - and screw it clockwise into the steering wheel top.

(4) Attach the steering wheel to the steering column

a) Unscrew the steering wheel bottom until the hole is wide enough to fit over the steering wheel column.

b) Put the steering wheel on the steering wheel column and tighten it (counter-clockwise).

(5) Attach the back-axles

a) Turn the chassis so that the steering wheel is on the bottom.

b) Slide the square peg of the back axle into the last notch on the chassis.

c) Slide the square peg of the other back axle into the opposite notch.

(6) Attach the towing piece

a) Turn the car so that the steering wheel is up.

b) Take the towing piece, and slide the square peg into the three-quarter hole at the back of the chassis. (The towing piece should be flat with the surface of the chassis.)

(7) Attach the wheels

a) One by one, stick the wheels on the axles so that the flat surface faces out, not in toward the car.

Writer 16

The following are instructions for the assembly of a Fischer Flat Bed Truck. Before you are a number of pieces of the kit that are not necessarily relevant for the assembly of this particular truck, as other objects or toys may be constructed from the objects in the kit. When fully assembled, the Flat Bed Truck will have four wheels stemming from a flat bed with a steering column and steering wheel.

To assemble the truck the steps are as follows:

1. Locate the Bed - The Bed is a red waffle-shaped item, the longest, largest and thickest item of the pieces before you and will be the basis of the truck assembly.

2. To the Truck Bed - Attach the red trailer hitch to the slot provided at either end of the Bed. The trailer hitch is the small bell-shaped item with two rectangular holes and connecting pin which should easily fit into the slot provided on the bed. Be careful not to confuse this trailer hitch with the trailer-hitch piece with three rectangular holes, which is also provided.

3. Locate the front axle - The front axle is longer than it is wide and has two stems or pins with slots therein, on opposite sides, to which pins will eventually attach the wheels.

4. Place the front axle under the Bed, under the first complete hole on the Bed opposite the end of the Bed where you placed the trailer hitch in Step 2.

5. Locate the steering column - the steering column is the long gray column with no bends and with the cube-like structure at one end.

6. Through the holes as aligned in step 4, push the steering column to the point where the cube on the end of the steering column is slightly less than flush with the bottom of the front axle.

7. Locate the Steering Column Nut - the Steering Column Nut fits cleanly around the steering column and is the shape of a doughnut with a bite taken out.

8. Slide the Steering Column Nut down the steering column until it is flush with the top of the Bed.

9. Locate the two rear axle components - There are two rear axle components, and they are not like the front axle which is one piece. The rear axle components do however have the same stems or pins for the connection of wheels as does the front axle. In addition they resemble the square gun turrets of mobile army tanks. The rear axle components have square shaped pins which connect to the Bed and are located on top of the rear axle components.

10. On the end of the Bed nearest to the trailer hitch, insert each rear axle component in the last slot on either side of the Bed, so that the rear axle component is on the bottom of the Bed or the same side of the Bed as the front axle.

11. Locate the Flat Bed Truck's wheels - There are four wheels and they are the only wheel-shaped objects of which there are four.

12. Through the holes provided in the wheels attach the wheels to the stems or pins in the front axle and rear axle components. The wheels should be attached so that the outside hub of the wheel away from the Bed is solid and not hollow.

13. Locate the parts for the steering wheel - The steering wheel consists of three parts - the Top, the Bottom, and Surrounding Circular Rubber Grip. The Top is one of four remaining identical pieces, is flat on top, has a hole through the center and has screws or threads leading from the bottom. The Bottom of the Steering Column is one of two identical remaining pieces and the only piece to which the threads or screws of the Top the steering wheel will attach. The Surrounding Circular Rubber Grip is black and rubber and is the smallest and smoothest of the three black rubber items in the Kit.

14. Place the Top of the Steering Wheel through the Surrounding Circular Rubber Grip so that the outside circumference of the top of the steering wheel is flush to the inside circumference of the Surrounding Circular Rubber Grip.

15. Pick up the Bottom of the Steering Wheel - The Bottom of the Steering Wheel should then be inserted into the thread provided by the Top of the Steering Wheel. The Bottom of the Steering Wheel should be turned clockwise to tighten the then assembled steering wheel structure.

16. Having assembled the Steering Wheel - Through the hole in the center of the Steering Wheel, insert the Steering Wheel Column, already a part of the Truck. Be sure to keep the bottom of the Steering Wheel facing the top of the Bed when doing this.

Writer 3

Assembly of Fischer Kit Vehicle

1. Selecting pieces.

   Select the following pieces necessary to assemble the vehicle:

   1 large red flat rectangular body with seven holes down the center

   4 red smooth wheels

   2 red square axles with cylinders

   1 red slender rectangular axle

   1 red flat oblong shaped piece with two openings

   1 red circular piece with nut

   1 red circular piece with bolt

   1 red small C-shaped piece

   1 black smooth rim

   1 straight gray cylindrical pin

2. Assemble rear wheels.

   a. Take the large rectangular body, the two red axles and two red wheels.

   b. Place the cylinder of the axle into the hole of the open end of the wheel. Repeat for other wheel.

   c. Insert raised square of one axle into the last open slot on body. Insert other axle into the opposite slot.

3. Assemble tow bar.

   Insert raised square of oblong piece into semicircle between the two rear wheels.

4. Assemble front wheels.

   a. Place wheels on the front axle by inserting both cylinders on the axle into the open ends of the two wheels.

   b. Attach the axle to the body.

      1) Insert the cylinder of the gray pin through the center square of the axle.

      2) Lock the square of the pin into the square of the axle.

      3) Insert the gray pin through bottom of the front hole of the red body piece.

5. Assemble the steering wheel.

   a. Snap the red bolt piece tightly into the black rim.

   b. Insert red nut piece with nut side down through the gray pin.

   c. Push the red bolt through gray pin.

   d. Screw nut tightly into bolt. Steering wheel should be at top of pin.

   e. Snap red C-shaped piece onto gray pin and push down tight against red body.

(Ring bell when finished.)

Writer 13

1. Put your collar up; if it's a button down collar, unbutton the collar then turn it up.

2. Second, put the tie around your neck with the thin end on your left side and the fat end on your right. Be sure to put the seamed part of the tie against the upturned collar, place the unseamed part of the tie along your torso.

3. Third, measure the tie along your torso; the fat end is approx. nine inches longer than the thin end.

4. Fourth, wrap the fat end around the thin end twice making sure that when you've finished wrapping, the pattern side of the tie faces forward.

5. Fifth, bring the fat end of the tie up through the opening at the throat making sure to bring the fat end through the back of the opening at the throat. Now, hold on to the fat end of the tie and push it through the second loop created in step four, that is, the second loop around the thin end of the tie.

6. Sixth, once you've put the fat end of the tie through the last loop, pull it taut so that it creates a knot.

7. Seventh, adjust the tie's length by holding the taut knot and pulling on the thin end of the tie. This will move the knot up to the throat and bring it to its correct position. A correctly-tied knot will leave a fat end that reaches the belt line without the thin end of the tie sticking out from underneath the fat part. If the thin part does stick out, you have to start again, this time beginning with a longer fat end than the first time.

Writer 4

Place the tie around your neck with the seam towards your body. The tie has one end which is much wider than the other. Slide the tie around your neck so that the wide end hangs 10-12 inches below the narrow end. Hold the wide end in your right hand and the narrow end in your left. Cross the wide end over the narrow end. While still holding the narrow end with your left hand, pull the wide end underneath with your right hand. Now the seam of the wide end will face away from you. Cross the wide end over the narrow end again with your right hand. You now have a closed loop of the tie around your neck. Now it is easier to switch hands: hold the narrow end with your right hand, and the wide end with your left. Push the wide end of the tie through the loop which is around your neck, and pull the wide end all the way through. Now, taking the wide end in your right hand again, lift it up and you will notice that there is now a horizontal loop of the tie around itself. Insert the wide end of the tie into the space directly behind the horizontal part. It will help if you hold the horizontal part with your left hand. Pull the wide end all the way through; as you pull this part down, the knot will tighten and get smaller. You do not need to make this knot very tight. Grasp the knot with your left hand, and the narrow part of the tie with your right. Pull down on the narrow part, which will move the knot closer to your neck.

If you have successfully made the knot, but the narrow end of the tie is hanging down below the wide end, you must retie the tie. By holding the knot with your right hand and pulling on the tie on the left side near your neck, you will undo the knot.

If you are going to retie the tie, start the wide end lower down than you did before.

Writer 12

How to Tie a Tie!

First put the tie under collar with under surface of tie facing against neck. Now position tie so that the thin end comes down to the fourth button (counting collar button). Position is done by pulling on fat end or thin end of tie until it is in proper place. Be sure that collar is completely covering tie.

Take section with fat end in right hand with thumb behind tie, cross this section over thin end, use left hand to hold the two sections together where they cross, thumb behind tie. This will be referred to as "the crossing." Grasp fat section of tie with right hand (thumb behind tie) about where thin section ends. Bring right hand up behind crossing and then move right hand in front of and just past crossing, keeping outer surface of tie facing forward. Place index finger of left hand in front of crossing so that thumb and index finger are pushing against each other. Release tie from right hand, then grasp tie with thumb and index finger of right hand at crossing so that tips of thumb and index finger of right and left hands are touching. Release left hand from tie. With left hand put fat section of tie behind and through the V formed above crossing, pushing section from behind through V, with thumb grasping that section of tie through the V between index and middle fingers. Then move left thumb from behind crossing and with left hand pull remainder of fat section completely out and in front of crossing. Grasp bottom end of fat section with left hand (thumb behind tie). Raise left hand above crossing, outer surface of tie facing forward. Remove right index finger from front of crossing, place it inside of outer loop of crossing against thumb of right hand. With left hand push tie between right index finger and outer loop. When edge of tie emerges beneath crossing release it from left hand, grasp protruding section of tie with left hand and pull remainder through. Hold bottom of this section with right hand, hold knot between thumb and first two fingers of left hand. Pull down with right hand--push knot up with left hand, simultaneously moving left hand toward the right to center the knot. Stop when knot seems snug against neck, but not too tight.

APPENDIX B

A Study of Revising Instructions

APPENDIX B

A Study of Revising Instructions

Revised assembly instructions. The eight revised sets of instructions differed less from one another than the eight original sets. In the revised sets, only the error measure was significant ($F[7,65]=2.48$, $p=.0253$), compared with error, functionality, and time in the original instructions. Table B-1 shows the mean errors made by users of each revised set of instructions, compared with the mean errors made by users of each original set.

Focusing on errors, we found that:

- The two least effective original instructions improved to about the level of the moderately effective original instructions.
- Of the five moderately effective original instructions, three improved to about the level of instructions of the most effective original instructions.
- One of the moderately effective instructions got worse (Writer 16).

Table B.1. Mean Total Errors for Users of Original and Revised Assembly Instructions

WRITER #	ORIGINAL INSTRUCTIONS	REVISED INSTRUCTIONS
15	2.0	1.8
16	3.1	4.8
10	3.0	1.8
6	3.8	1.7
2	4.2	2.6
7	5.0	5.2
9	6.5	3.9
3	6.5	4.1

- One of the moderately effective instructions stayed at the same level (Writer 7).

- The most effective instructions get slightly better (reaching asymptote on functionality, producing somewhat fewer mean errors).

The data on functionality mirror the error patterns.

Thus, the predominant effect of revising with feedback was to create modest increases in effectiveness. The reason for modest increases (no more and no less) seems to be that writers focused specifically on the errors they saw the user making in the videotaped feedback session. The typical change was to add information necessary to specify a task dimension that the writer originally failed to specify: e.g., adding orientation information when the user attached a piece upside down. Writers rarely changed other information.

The reason for lack of improvement in one writer's instructions (Writer 7) seems to be that he was an ineffective reviser. He detected the errors users make and found the relevant points in his instructions, but his corrections did not seem to reduce ambiguities.

The reason for lack of improvement in the other writer's instructions (Writer 16) seems to be a function of the kind of user he saw. This user made one type of common error but not another type of common error. The writer corrected only the ambiguities related to the first type of error. Users of the

revised instructions made more of the second type of error than did users of the original instructions.

Looking at change processes, we found that all writers made changes to their instructions, both before and after they saw the videotapes. They made more task-related changes after feedback than before feedback. Moreover, in revising with the videotapes, all writers addressed some informational ambiguities that they had not detected in revising before they saw the videotape.

Revised tie-tying instructions. The eight revised sets of instructions did not differ significantly from each other on the accuracy measure ($\underline{F}[7,46]=1.397$, $p = .2298$). The range of mean correctness scores was narrower (1.1-1.9) than on the original instructions (.5-1.7) and the end points were higher. Table B-2 shows the mean correctness scores of users of each revised set of instructions, compared with the correctness scores of users of each original set.

Looking at individual instructions, we found that:

- Four sets of instructions stayed at approximately the same level between original and revised: Writers 11, 13, and 8 (effective) and Writer 14 (ineffective).

- Three sets of instructions improved: Writer 5 (effective) and Writers 4 and 12 (ineffective). The improvement in Writer 4 was dramatic.

TABLE B.2. Mean Total Correctiness Scores for Users of Original and Revised Tie-Tying Instructions

WRITER #	ORIGINAL INSTRUCTIONS	REVISED INSTRUCTIONS
11	1.7	1.8
13	1.6	1.5
5	1.6	1.9
8	1.5	1.5
1	1.5	1.1
14	1.0	1.1
4	1.0	1.9
12	.5	1.3

- One set of instructions got worse: Writer 1 went from a level of moderately effective to a level of the ineffective (based on our original cut-off points).

The decline in Writer 1's instructions is attributable to the fact that he inadvertently eliminated a step in making extensive changes to his draft while viewing the videotape. He did not review his draft again after making changes. This result confirms the obvious: A whole-draft review <u>after</u> the writer makes extensive changes is important to be sure essential information hasn't been changed.

The lack of increase in four of the writers can be attributed either to the fact that the writer had already achieved effective instructions (Writer 11) or to the fact that the writer viewed a videotape that did not reveal all the ambiguities in the instructions. Unlike assembly, it is impossible for a user of the tie-tying instructions to make more than one critical error, since all task operations are serially dependent. Thus, unless the user on the videotape backtracked or started over, and then made new errors, the writer could theoretically infer only one source of ambiguity in the instructions.

One of the least effective instructions improved dramatically. This seems to be because the instructions were most ambiguous at one critical point, a point captured by the

user's errors and comments and subsequently addressed by the writer.

Writers' change processes corresponded to those of the assembly writers. One tie-tying writer, however (ineffective Writer 14), was able to detect the ambiguity in one of his over-segmented descriptions of a task step before he saw the videotape. He corrected the ambiguity at that point and changed little else while viewing the videotape.

Summary. In general, we infer that seeing an instruction user make errors can help writers pinpoint ambiguities in their instructions. However, the ambiguities writers detect appear to depend on the particular user they see. A single user may not make all the types of errors that are relatively common across users. Theoretically and practically, it would be of interest to see how writers revise based on a written list of the most common errors made by a sample of users: Would a comprehensive list of errors be more effective than an observation of performance in helping writers detect ambiguities in their instructions? Alternatively, are there aspects peculiar to observing that aid detection or diagnosis more than an error summary would?

Navy

1 Robert Ahlers
Code N711
Human Factors Laboratory
NAVTRAEQUIPCEN
Orlando, FL 32813

1 Dr. Ed Aiken
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Meryl S. Baker
Navy Personnel R&D Center
San Diego, CA 92152

1 Code N711
Attn: Arthur S. Blaiwes
Naval Training Equipment Center
Orlando, FL 32813

1 Dr. Richard Braby
TAEG
Naval Training Equipment Center
Orlando, FL 32813

1 Dr. Robert Breaux
NAVTRAEQUIPCEN
Code N-095R
Orlando, FL 32813

1 Dr. Richard Cantone
Navy Research Laboratory
Code 7510
Washington, DC 20375

1 Dr. Robert Carroll
NAVOP 115
Washington , DC 20370

1 Dr. Fred Chang
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Susan Chipman
Code 442PT
Office of Naval Research
800 N. Quincy St.
Arlington, VA 22217

1 Dr. Stanley Collyer
Office of Naval Technology
800 N. Quincy Street
Arlington, VA 22217

Navy

1 CDR Mike Curran
Office of Naval Research
800 N. Quincy St.
Code 270
Arlington, VA 22217

1 Dr. Charles E. Davis
Personnel and Training Research
Office of Naval Research (Code 442PT)
800 North Quincy Street
Arlington, VA 22217

1 Edward E. Eddowes
CNATRA N301
Naval Air Station
Corpus Christi, TX 78419

1 Dr. Marshall J. Farr
Director, Psychological Sciences
Office of Naval Research (Code 442)
800 North Quincy Street
Arlington, VA 22217

1 Dr. Jude Franklin
Code 7510
Navy Research Laboratory
Washington, DC 20375

1 Dr. Jim Hollan
Code 14
Navy Personnel R & D Center
San Diego, CA 92152

1 Dr. Ed Hutchins
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Norman J. Kerr
Chief of Naval Technical Training
Naval Air Station Memphis (75)
Millington, TN 38054

1 Dr. Peter Kincaid
Training Analysis & Evaluation Group
Dept. of the Navy
Orlando, FL 32813

1 Dr. William L. Maloy (02)
Chief of Naval Education and Training
Naval Air Station
Pensacola, FL 32508

Navy

1 Dr. Joe McLachlan
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. James McMichael
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr William Montague
NPRDC Code 13
San Diego, CA 92152

1 Office of Naval Research
Code 433
800 N. Quincy SStreet
Arlington, VA 22217

1 Director
Engineering Psychology Group
Code 442EP
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

1 Organizational Effectiveness
Research Group, Code 442OE
Office of Naval Research
Arlington, VA 22217

1 Psychologist
ONR Branch Office
1030 East Green Street
Pasadena, CA 91101

1 LT Frank C. Petho, MSC, USN (Ph.D)
CNET (N-432)
NAS
Pensacola, FL 32508

1 Dr. Gil Ricard
Code N711
NTEC
Orlando, FL 32813

1 Dr. Paul J. Schneck
Office of Naval Research
Code 433
800 N. Quincy
Arlington, VA 22217

1 Dr. Michael G. Shafto
ONR Code 442PT
800 N. Quincy Street
Arlington, VA 22217

Navy

1 Dr. Douglas Smith
Office of Naval Research
Code 433
800 N. Quincy
Arlington, VA 22217

1 Dr. Robert G. Smith
Office of Chief of Naval Operations
OP-987H
Washington, DC 20350

1 Dr. Alfred F. Smode
Senior Scientist
Code 7B
Naval Training Equipment Center
Orlando, FL 32813

1 Dr. Richard Sorensen
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Frederick Steinheiser
CNO - OP115
Navy Annex
Arlington, VA 20370

1 Dr. Thomas Sticht
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Martin A. Tolcott
Leader, Psychological Sciences Division
Office of Naval Research
800 N. Quincy St.
Arlinsgon, VA 22217

1 Dr. James Tweeddale
Technical Director
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Nick Van Matre
CNET
NAS
Pensacola, FL 32508

1 Dr. Wallace Wulfeck, III
Navy Personnel R&D Center
San Diego, CA 92152

Navy

1 Dr. Steven Zornetzer
Associate Director for Life Sciences
Office of Naval Research
800 N. Quincy St.
Arlington, VA 22217

Army

1 Technical Director
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Beatrice J. Farr
U. S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Milton S. Katz
Training Technical Area
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Harold F. O'Neil, Jr.
Director, Training Research Lab
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Commander, U.S. Army Research Institute
for the Behavioral & Social Sciences
ATTN: PERI-BR (Dr. Judith Orasanu)
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Joseph Psotka, Ph.D.
ATTN: PERI-1C
Army Research Institute
5001 Eisenhower Ave.
Alexandria, VA 22333

1 Dr. Robert Sasmor
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 DR. ROBERT J. SEIDEL
US Army Research Institute
5001 Eisenhower Ave.
300 N. WASHINGTON ST.
Alexandria, VA 22333

Air Force

1 U.S. Air Force Office of Scientific
Research
Life Sciences Directorate, NL
Bolling Air Force Base
Washington, DC 20332

1 Dr. Earl A. Alluisi
HQ, AFHRL (AFSC)
Brooks AFB, TX 78235

1 Bryan Dallman
AFHRL/LRT
Lowry AFB, CO 80230

1 Dr. Genevieve Haddad
Program Manager
Life Sciences Directorate
AFOSR
Bolling AFB, DC 20332

1 Dr. T. M. Longridge
AFHRL/OTE
Williams AFB, AZ 85224

1 Dr. John Tangney
AFOSR/NL
Bolling AFB, DC 20332

1 Dr. Joseph Yasatuke
AFHRL/LRT
Lowry AFB, CO 80230

Department of Defense

1 Dr. Craig I. Fields
Advanced Research Projects Agency
1400 Wilson Blvd.
Arlington, VA 22209

1 Military Assistant for Training and
Personnel Technology
Office of the Under Secretary of Defens
for Research & Engineering
Room 3D129, The Pentagon
Washington, DC 20301

1 Major Jack Thorpe
DARPA
1400 Wilson Blvd.
Arlington, VA 22209

1 Dr. Robert A. Wisher
OUSDRE (ELS)
The Pentagon, Room 3D129
Washington, DC 20301

Civilian Agencies

1 Edward Esty
Department of Education, OERI
MS 40
1200 19th St., NW
Washington, DC 20208

1 Dr. Arthur Melmed
724 Brown
U. S. Dept. of Education
Washington, DC 20208

1 Dr. Andrew R. Molnar
Office of Scientific and Engineering
Personnel and Education
National Science Foundation
Washington, DC 20550

1 Dr. Ramsay W. Selden
National Institute of Education
1200 19th St., NW
Washington, DC 20208

1 Dr. Edward C. Weiss
National Science Foundation
1800 G Street, NW
Washington, DC 20550

1 Dr. Frank Withrow
U. S. Office of Education
400 Maryland Ave. SW
Washington, DC 20202

1 Dr. Joseph L. Young, Director
Memory & Cognitive Processes
National Science Foundation
Washington, DC 20550

Private Sector

1 Dr. John R. Anderson
Department of Psychology
Carnegie-Mellon University
Pittsburgh, PA 15213

1 Patricia Baggett
Department of Psychology
University of Colorado
Boulder, CO 80309

1 Eva L. Baker
Director
UCLA Center for the Study of Evaluation
145 Moore Hall
University of California, Los Angeles
Los Angeles, CA 90024

1 Mr. Avron Barr
Department of Computer Science
Stanford University
Stanford, CA 94305

1 Dr. John Black
Yale University
Box 11A, Yale Station
New Haven, CT 06520

1 Dr. John S. Brown
XEROX Palo Alto Research Center
3333 Coyote Road
Palo Alto, CA 94304

1 Dr. Bruce Buchanan
Department of Computer Science
Stanford University
Stanford, CA 94305

1 Dr. Jaime Carbonell
Carnegie-Mellon University
Department of Psychology
Pittsburgh, PA 15213

1 Dr. Pat Carpenter
Department of Psychology
Carnegie-Mellon University
Pittsburgh, PA 15213

1 Dr. Davida Charney
Department of Psychology
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Private Sector

1 Eugene Charniak
Department of Computer Science
Brown University
Providence, RI 02912

1 Dr. Micheline Chi
Learning R & D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15213

1 Dr. William Clancey
Department of Computer Science
Stanford University
Stanford, CA 94306

1 Dr. Michael Cole
University of California
at San Diego
Laboratory of Comparative
Human Cognition - D003A
La Jolla, CA 92093

1 Dr. Allan M. Collins
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, MA 02138

1 Dr. Emmanuel Donchin
Department of Psychology
University of Illinois
Champaign, IL 61820

1 Dr. Thomas M. Duffy
Department of English
Carnegie-Mellon University
Schenley Park
Pittsburgh, CA 15213

1 Dr. Anders Ericsson
Department of Psychology
University of Colorado
Boulder, CO 30309

1 Dr. Paul Feltovich
Department of Medical Education
Southern Illinois University
School of Medicine
P.O. Box 3926
Springfield, IL 62708

Private Sector

1 Mr. Wallace Feurzeig
Department of Educational Technology
Bolt Beranek & Newman
10 Moulton St.
Cambridge, MA 02238

1 Dr. Dexter Fletcher
University of Oregon
Department of Computer Science
Eugene, OR 97403

1 Dr. John R. Frederiksen
Bolt Beranek & Newman
50 Moulton Street
Cambridge, MA 02138

1 Dr. Michael Genesereth
Department of Computer Science
Stanford University
Stanford, CA 94305

1 Dr. Dedre Gentner
Bolt Beranek & Newman
10 Moulton St.
Cambridge, MA 02138

1 Dr. Robert Glaser
Learning Research & Development Center
University of Pittsburgh
3939 O'Hara Street
PITTSBURGH, PA 15260

1 Dr. Marvin D. Glock
217 Stone Hall
Cornell University
Ithaca, NY 14853

1 Dr. Josph Goguen
SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025

1 DR. JAMES G. GREENO
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

1 Dr. Henry M. Halff
Halff Resources
4918 33rd Road, North
Arlington, VA 22207

Private Sector

1 Dr. Reid Hastie
Department of Psychology
Northwestern University
Evanston, IL 60201

1 Dr. Barbara Hayes-Roth
Department of Computer Science
Stanford University
Stanford, CA 95305

1 Dr. Joan I. Heller
Graduate Group in Science and
Mathematics Education
c/o School of Education
University of California
Berkeley, CA 94720

1 Dr. James R. Hoffman
Department of Psychology
University of Delaware
Newark, DE 19711

1 Melissa Holland
American Institutes for Research
1055 Thomas Jefferson St., N.W.
Washington, DC 20007

1 Dr. Kristina Hooper
Corporate Research, ATARI
1196 Borregas
Sunnyvale, CA 94086

1 Dr. Marcel Just
Department of Psychology
Carnegie-Mellon University
Pittsburgh, PA 15213

1 Dr. David Kieras
Department of Psychology
University of Arizona
Tuscon, AZ 85721

1 Dr. Walter Kintsch
Department of Psychology
University of Colorado
Boulder, CO 80302

1 Dr. Pat Langley
The Robotics Institute
Carnegie-Mellon University
Pittsburgh, PA 15213

Private Sector

1 Dr. Jill Larkin
Departrent of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213

1 Dr. Alan Lesgold
Learning R&D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

1 Dr. Jim Lavin
University of California
at San Diego
Laboratory fof Comparative
Human Cognition - D003A
La Jolla, CA 92093

1 Dr. Marcia C. Linn
Lawrence Hall of Science
University of California
Berkeley, CA 94720

1 Dr. Don Lyon
P. O. Box 44
Higley , AZ 85236

1 Dr. Jay McClelland
Departrent of Psychology
MIT
Cambridge, MA 02139

1 Dr. James R. Miller
Computer*Thought Corporation
1721 West Plano Highway
Plano, TX 75075

1 Dr. Allen Munro
Behavioral Technology Laboratories
1845 Elena Ave., Fourth Floor
Redondo Beach, CA 90277

1 Dr. Donald A Norman
Cognitive Science, C-015
Univ. of California, San Diego
La Jolla, CA 92093

1 Dr. Jesse Orlansky
Institute for Defense Analyses
1801 N. Beauregard St.
Alexandria, VA 22311

Private Sector

1 Dr. Nancy Pennington
University of Chicago
Graduate School of Business
1101 E. 58th St.
Chicago, IL 60637

1 Dr. Ann Piestrup
THE LEARNING COMPANY
545 Middlefield Road, Suite 170
Menlo Park, CA 94025

1 Dr. Steven E. Poltrock
Bell Laboratories 2D-444
600 Mountain Ave.
Murray Hill, NJ 07974

1 Dr. Lynne Reder
Department of Psychology
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

1 Dr. Fred Reif
Physics Department
University of California
Berkeley, CA 94720

1 Dr. Lauren Resnick
LRDC
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 1521

1 Dr. Jeff Richardson
Denver Research Institute
University of Denver
Denver, CO 80208

1 Mary S. Riley
Program in Cognitive Science
Center for Human Information Processing
University of California, San Diego
La Jolla, CA 92093

1 Dr. Andrew M. Rose
American Institutes for Research
1055 Thomas Jefferson St. NW
Washington, DC 20007

1 Dr. William B. Rouse
Georgia Institute of Technology
School of Industrial & Systems
Engineering
Atlanta, GA 30332

Private Sector

1 Dr. Roger Schank
Yale University
Department of Computer Science
P.O. Box 2158
New Haven, CT 06520

1 Dr. Walter Schneider
Psychology Department
603 E. Daniel
Champaign, IL 61820

1 Dr. Alan Schoenfeld
Mathematics and Education
The University of Rochester
Rochester, NY 14627

1 Dr. Ted Shortliffe
Computer Science Department
Stanford University
Stanford, CA 94305

1 Dr. Edward E. Smith
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, MA 02138

1 Dr. Eliott Soloway
Yale University
Department of Computer Science
P.O. Box 2158
New Haven, CT 06520

1 Dr. Kathryn T. Spoehr
Psychology Department
Brown University
Providence, RI 02912

1 Dr. Robert Sternberg
Dept. of Psychology
Yale University
Box 11A, Yale Station
New Haven. CT 06520

1 Dr. Albert Stevens
Bolt Beranek & Newman, Inc.
10 Moulton St.
Cambridge, MA 02238

1 Dr. David Stone
KAJ Software, Inc.
3420 East Shea Blvd.
Suite 161
Pheonix, AZ 85028

Private Sector

1 DR. PATRICK SUPPES
INSTITUTE FOR MATHEMATICAL STUDIES IN
THE SOCIAL SCIENCES
STANFORD UNIVERSITY
STANFORD, CA 94305

1 Dr. Kikumi Tatsuoka
Computer Based Education Research Lab
252 Engineering Research Laboratory
Urbana, IL 61801

1 Dr. Perry W. Thorndyke
Perceptronics, Inc.
545 Middlefield Road, Suite 140
Menlo Park, CA 94025

1 Dr. Douglas Towne
Univ. of So. California
Behavioral Technology Labs
1845 S. Elena Ave.
Redondo Beach, CA 90277

1 Dr. Kurt Van Lehn
Xerox PARC
3333 Coyote Hill Road
Palo Alto, CA 94304

1 Beth Warren
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, MA 02138

1 Dr. Keith T. Wescourt
Perceptronics, Inc.
545 Middlefield Road, Suite 140
Menlo Park, CA 94025

1 Dr. Mike Williams
IntelliGenetics
124 University Avenue
Palo Alto, CA 94301

END

FILMED

11-85

DTIC